

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 10:44:58 ; Search time 11824 Seconds

(without alignments)
11385.611 Million cell updates/sec

Title: US-09-900-751-1

Perfect score: 3106

Sequence: 1 catggtacgacgctgcccgg.....ttaaaaaaaaaaaaaaa 3106

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3106	100.0	3106	6	AX395268
2	3106	100.0	3106	10	AF042822
3	3102.8	99.9	3248	10	BC005496
4	2595	83.5	3174	10	AB049189
5	2313.6	74.5	2568	10	AB037898
6	1905.6	61.4	3273	9	BC030532
7	1889.8	60.8	3149	9	BD274671
8	1884.8	60.7	3120	9	AF118224
9	1884.6	60.7	3128	9	AB030036
10	1883.2	60.6	3147	6	AR081724
11	1883.2	60.6	3147	6	AR229704
12	1883.2	60.6	3147	6	AR229712
13	1883.2	60.6	3147	6	AR430930
14	1883.2	60.6	3147	6	AR430938
15	1883.2	60.6	3147	6	AX207897
16	1883.2	60.6	3147	6	AX207945
17	1883.2	60.6	3147	6	AF057145
18	1883.2	60.6	3147	6	BD274670
19	1752.6	56.4	2955	6	AR081725
20	1530.8	49.3	2900	6	AR229705
21	1530.8	49.3	2900	6	AR430931
22	1530.8	49.3	2900	6	HSU20428
23	1530.8	49.3	2900	6	AX921703
24	1518.2	48.9	2838	6	BC005826
25	1175.8	37.9	2139	9	BC018146
26	945.4	30.4	1823	9	BC018146
27	943.8	30.4	2152	6	AR379612
28	809.6	26.1	3487	5	AB038498
29	701.2	22.6	1553	6	AR263832
30	623.2	20.1	235924	10	AC114542
31	556.4	17.9	726	6	AX473052
32	551.8	17.8	723	6	E13204
33	359.6	11.6	236387	2	AC096042
34	359.6	11.6	261906	2	AC112632
35	359.6	11.6	277797	2	AC128347
36	306.8	9.9	434	6	AR273737
37	306.8	9.9	434	6	AR273738
38	306.8	9.9	434	6	AR407593
39	306.8	9.9	434	6	AX368770
40	265	8.5	591	6	AX261852
41	244.8	7.9	3142	10	BC057674
42	244.8	7.9	3183	10	AY240929
43	244.8	7.9	3183	10	BC029645
44	243.6	7.8	1485	6	AX814628
45	242.4	7.8	3194	9	AY055384

ALIGNMENTS

```

RESULT 1
LOCUS AX395268 3106 bp DNA linear PAT 18-MAY-2002
DEFINITION Sequence 5 from Patent WO0203787.
ACCESSION AX395268
VERSION AX395268.1 GI:21066293
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Allen, K.D. and Leviten, M.W.
Transgenic mice containing targeted gene disruptions
Patent: WO 0203787-A 5 17-JAN-2002;
JOURNAL

```

FEATURES
 source
 Delagen, Inc. (US)
 Location/Qualifiers
 1..3106
 /organism="Mus musculus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10090"

ORIGIN

Query Match 100.0%; Score 3106; DB 6; Length 3106;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATGTGAGAGGCTGCCCCGAGGACCAAGCTCTGAGACCGGATCGGACCGCAAA 60
 1 CATGTGAGAGGCTGCCCCGAGGACCAAGCTCTGAGACCGGATCGGACCGCAAA 60
 61 CCATGGGTAGCAATCGGGGCGGCAAGGCGGAGGGGCTCTCAGGACTTCGGGCGGGAC 120
 61 CCATGGGTAGCAATCGGGGCGGCAAGGCGGAGGGGCTCTCAGGACTTCGGGCGGGAC 120
 121 TCAAGTACAACTCCGGGCTAGAGAACTGAATGCTTTGAGAGAGGTGTGAGTCTTCC 180
 121 TCAAGTACAACTCCGGGCTAGAGAACTGAATGCTTTGAGAGAGGTGTGAGTCTTCC 180
 181 CTGGGAACATGCCAAGAAAGTGGAGAAAGGAGGCCCGGAGCGCTGGGTGGTGGTGG 240
 181 CTGGGAACATGCCAAGAAAGTGGAGAAAGGAGGCCCGGAGCGCTGGGTGGTGGTGG 240
 241 CAGTGTCTGAGCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
 241 CAGTGTCTGAGCTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
 301 ATTATCGAATGTGCGGGTTCAAAAAGTCTTCAATGCGCATCTGAGATCACAATAGA 360
 301 ATTATCGAATGTGCGGGTTCAAAAAGTCTTCAATGCGCATCTGAGATCACAATAGA 360
 361 TCTTCTGAGAGGCTAGAGAACTCACTCCAGAGTTTATGAGCTTGGCCAGCGAGG 420
 361 TCTTCTGAGAGGCTAGAGAACTCACTCCAGAGTTTATGAGCTTGGCCAGCGAGG 420
 421 TGAAGAGGCGCTGAGAGCTCAATGAATGCCCTGTCTGGGTCCTTACCAAGA 480
 421 TGAAGAGGCGCTGAGAGCTCAATGAATGCCCTGTCTGGGTCCTTACCAAGA 480
 481 AGTGGCTGTACTGCTTCTGAGAGGAGCTGTATGCTCTTACTGAGAGTTC 540
 481 AGTGGCTGTACTGCTTCTGAGAGGAGCTGTATGCTCTTACTGAGAGTTC 540
 541 GATGCGGCTGTACTGCTTCTGAGAGGAGCTGTATGCTCTTACTGAGAGTTC 540
 541 GATGCGGCTGTACTGCTTCTGAGAGGAGCTGTATGCTCTTACTGAGAGTTC 540
 601 CATTGCGACCCCGAGACGAGGCACTGAATCTTGTGCTTAACTCTGTGTGCTTTC 660
 601 CATTGCGACCCCGAGACGAGGCACTGAATCTTGTGCTTAACTCTGTGTGCTTTC 660
 661 CATGCGACCCCGAGACGAGGCACTGAATCTTGTGCTTAACTCTGTGTGCTTTC 660
 661 CATGCGACCCCGAGACGAGGCACTGAATCTTGTGCTTAACTCTGTGTGCTTTC 660
 721 CCATGAGCCCGAGAGTGTGAGAGGAGCTCAGAGAAAGCTGTGAGTTCCTGAGAG 720
 721 CCATGAGCCCGAGAGTGTGAGAGGAGCTCAGAGAAAGCTGTGAGTTCCTGAGAG 720
 781 CCATGAGCCCGAGAGTGTGAGAGGAGCTCAGAGAAAGCTGTGAGTTCCTGAGAG 780
 781 CCATGAGCCCGAGAGTGTGAGAGGAGCTCAGAGAAAGCTGTGAGTTCCTGAGAG 780
 841 TCCGAGAGCTTGTATGTGCTCTCTCTGTGAGAGCATGTGAGAGCTCTGAGAGCT 840
 841 TCCGAGAGCTTGTATGTGCTCTCTCTGTGAGAGCATGTGAGAGCTCTGAGAGCT 840
 901 ATAGGCTGAGCCCATGGAACCCACGCTGTGTGCGGCTGTGTGACCTTCTCAACCT 960

901 ATAGGCTGAGCCCATGGAACCCACGCTGTGTGCGGCTGTGTGACCTTCTCAACCT 960
 961 CCTACCACTGATTTCTTCTCTCTCCCAAGAGCTCTTCTGAGAGCTATACCAATA 1020
 961 CCTACCACTGATTTCTTCTCTCTCCCAAGAGCTCTTCTGAGAGCTATACCAATA 1020
 1021 CTGACCGGCGCATCTCTGCTTGAAGGCACTTCTTCCAGTGTGCCAAGATGAGCACT 1080
 1021 CTGACCGGCGCATCTCTGCTTGAAGGCACTTCTTCCAGTGTGCCAAGATGAGCACT 1080
 1081 GTGGCGGCTTTTGAAGTGAACCCCAAGGAGCATTTGAGAGGCTTACTATCCAGGCT 1140
 1081 GTGGCGGCTTTTGAAGTGAACCCCAAGGAGCATTTGAGAGGCTTACTATCCAGGCT 1140
 1141 ACCCGGCCCAATCACTGAGCAATGAAATTCAAAGTGCCTCAACCGGAGCGTGAAG 1200
 1141 ACCCGGCCCAATCACTGAGCAATGAAATTCAAAGTGCCTCAACCGGAGCGTGAAG 1200
 1201 TGGGCTTCAAACTCTTCTATCTGAGTGAACCCCAAGTACCAAGTGTCTCTGACCAAG 1260
 1201 TGGGCTTCAAACTCTTCTATCTGAGTGAACCCCAAGTACCAAGTGTCTCTGACCAAG 1260
 1261 ACTATGTGAGATCAACGGGAGAGATCTGCGGTGAGAGGCTCCAGTTGTGTGAGCA 1320
 1261 ACTATGTGAGATCAACGGGAGAGATCTGCGGTGAGAGGCTCCAGTTGTGTGAGCA 1320
 1321 GCAACACGACCAAGATTAAGTCACTTCCATTCATCTGATCACTGATACCGGAGT 1380
 1321 GCAACACGACCAAGATTAAGTCACTTCCATTCATCTGATCACTGATACCGGAGT 1380
 1381 TCTAGCTGATGATCTCTCTCTCAAGCTCAACGACCGGTGCCAGGATGTTCAATGCA 1440
 1381 TCTAGCTGATGATCTCTCTCTCAAGCTCAACGACCGGTGCCAGGATGTTCAATGCA 1440
 1441 AGACTGACGCTGATCCGAAGAGATCTGCGGTGAGAGGCTGAGAGCTGAGAGT 1500
 1441 AGACTGACGCTGATCCGAAGAGATCTGCGGTGAGAGGCTGAGAGCTGAGAGT 1500
 1501 AGACTGACGCTGATCCGAAGAGATCTGCGGTGAGAGGCTGAGAGCTGAGAGT 1560
 1501 AGACTGACGCTGATCCGAAGAGATCTGCGGTGAGAGGCTGAGAGCTGAGAGT 1560
 1561 TCTGCAAGCCCTCTTCTGAGTGTGAGCAAGTGTCAACGATGTGAGAGCGAAGTGAAG 1620
 1561 TCTGCAAGCCCTCTTCTGAGTGTGAGCAAGTGTCAACGATGTGAGAGCGAAGTGAAG 1620
 1621 AGAGGCTGAGAGCTGCTCTCTGAGAGGATTTCAAGTGTCAATGAGAGAGTCTCTC 1680
 1621 AGAGGCTGAGAGCTGCTCTCTGAGAGGATTTCAAGTGTCAATGAGAGAGTCTCTC 1680
 1681 AGAGGCTGAGAGCTGCTCTCTGAGAGGATTTCAAGTGTCAATGAGAGAGTCTCTC 1740
 1681 AGAGGCTGAGAGCTGCTCTCTGAGAGGATTTCAAGTGTCAATGAGAGAGTCTCTC 1740
 1741 AAGGCTGAGAGTGTGCTCTCTGAGAGGATTTCAAGTGTCAATGAGAGAGTCTCTC 1800
 1741 AAGGCTGAGAGTGTGCTCTCTGAGAGGATTTCAAGTGTCAATGAGAGAGTCTCTC 1800
 1801 TGAAGAGGCGCAACCTGAGAGTGAATGAGAGAGGAGCTGAGAGAGTGTGAGAG 1860
 1801 TGAAGAGGCGCAACCTGAGAGTGAATGAGAGAGGAGCTGAGAGAGTGTGAGAG 1860
 1861 TGAAGAGGCGCAACCTGAGAGTGAATGAGAGAGGAGCTGAGAGAGTGTGAGAG 1920
 1861 TGAAGAGGCGCAACCTGAGAGTGAATGAGAGAGGAGCTGAGAGAGTGTGAGAG 1920
 1921 ATGCGAGAGAGGCGAGTGGGCTTCTGAGAGGAGCTTCAAGGCTGAGAGGCTGAGAG 1980
 1921 ATGCGAGAGAGGCGAGTGGGCTTCTGAGAGGAGCTTCAAGGCTGAGAGGCTGAGAG 1980
 1981 TGTGTGGGCTCTGCTCATCTCTCTGAGTGTGCTGAGTGTGAGTGTGAGTGTGAG 2040

```

Db      1981  TGTGTGGGGCGCTGCTCATCTCTCTGACTGGCTGTGTCTCTGCAAGCTCATTTGCTTTAGG  2040
Qy      2041  ATGCAAAAATTTCAAGTACTCAGACTACAGATGTGTGACGGCCCTTCTGCGTCTGCTGG  2100
Db      2041  ATGCAAAAATTTCAAGTACTCAGACTACAGATGTGTGACGGCCCTTCTGCGTCTGCTGG  2100
Qy      2101  ACCGAGCAAGCGCAGGCGCTGTGGGGTGAAGAGCTGAAGCTCAACGATCATCAACC  2160
Db      2101  ACCGAGCAAGCGCAGGCGCTGTGGGGTGAAGAGCTGAAGCTCAACGATCATCAACC  2160
Qy      2161  ACCCTTCTTCAATGATTTGATCTGATGATGATGATGATGATGATGATGATGATGAT  2220
Db      2161  ACCCTTCTTCAATGATTTGATCTGATGATGATGATGATGATGATGATGATGATGAT  2220
Qy      2221  CGGTGAGTACAGACCGCTGTGTGCGCCCACTGCTGCTGATGATGATGATGATGATGAT  2280
Db      2221  CGGTGAGTACAGACCGCTGTGTGCGCCCACTGCTGCTGATGATGATGATGATGATGAT  2280
Qy      2281  CTGCTGGCAAGGCGCATCTGGGTGACAGGCTGGGGGCAACAAAAGAGGATACCGAG  2340
Db      2281  CTGCTGGCAAGGCGCATCTGGGTGACAGGCTGGGGGCAACAAAAGAGGATACCGAG  2340
Qy      2341  CGCTGATCTGCAAGAGGTGAGATCCGTGATCAACAGACCACTGATGAGACCTCA  2400
Db      2341  CGCTGATCTGCAAGAGGTGAGATCCGTGATCAACAGACCACTGATGAGACCTCA  2400
Qy      2401  TGCCTGAGAGATCAACCCCAAGATATGTGTGTGGGTTTCTCACTGGGGGTGTGACT  2460
Db      2401  TGCCTGAGAGATCAACCCCAAGATATGTGTGTGGGTTTCTCACTGGGGGTGTGACT  2460
Qy      2461  CCTGCAAGGATGATCTGTGTGGCCCTTGTCAAGCCCGGAGAAAGATGGCCGATTTTC  2520
Db      2461  CCTGCAAGGATGATCTGTGTGGCCCTTGTCAAGCCCGGAGAAAGATGGCCGATTTTC  2520
Qy      2521  AGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT  2580
Db      2521  AGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT  2580
Qy      2581  CAAAGCTCCCTGTATGTTGCGGACTGTGATTAAGAGACCACTGGGGTATAGCAGATGAC  2640
Db      2581  CAAAGCTCCCTGTATGTTGCGGACTGTGATTAAGAGACCACTGGGGTATAGCAGATGAC  2640
Qy      2641  AGAGACCGGACCAACCAACCAAGGATGCGGACCACTGATGATGATGATGATGATGATGAT  2700
Db      2641  AGAGACCGGACCAACCAACCAAGGATGCGGACCACTGATGATGATGATGATGATGATGAT  2700
Qy      2701  GAAACATGACGATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT  2760
Db      2701  GAAACATGACGATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT  2760
Qy      2761  GCATCTTATGAGACTCAGAGTTTTCAAAAGTGTGAGACCCCTTCAGAGTTGAGAGAAC  2820
Db      2761  GCATCTTATGAGACTCAGAGTTTTCAAAAGTGTGAGACCCCTTCAGAGTTGAGAGAAC  2820
Qy      2821  TTGCGTGTAGGCGGCGGCTGTGGGGGCAAGGCTTTATGAGAGCTTCCCTCTAGCC  2880
Db      2821  TTGCGTGTAGGCGGCGGCTGTGGGGGCAAGGCTTTATGAGAGCTTCCCTCTAGCC  2880
Qy      2881  CTGAGCTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  2940
Db      2881  CTGAGCTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  2940
Qy      2941  GGGAGCCCTATGAGAGGAGGCGCTCAGGGCTCATCTTTTCAAGAAAGCGCCAGCTAGGA  3000
Db      2941  GGGAGCCCTATGAGAGGAGGCGCTCAGGGCTCATCTTTTCAAGAAAGCGCCAGCTAGGA  3000
Qy      3001  ACCCAAGAAAAGAGTGTACCTTAAGCTGAATGTTTGTGTGTGTGTGTGTGTGTGTGTGT  3060
Db      3001  ACCCAAGAAAAGAGTGTACCTTAAGCTGAATGTTTGTGTGTGTGTGTGTGTGTGTGTGT  3060
Qy      3061  TTGAGAGTAAACAATTTATTTCTTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAA  3106
Db      3061  TTGAGAGTAAACAATTTATTTCTTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAA  3106

```

RESULT 2

AF042822

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

TITLE

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

ORIGIN

AKAIVTWGHTKXGNGALILKGEIRVINOTTCEDILPMQOITPRMVCVFLSGV
 DSCQDSGSPJLSAEKDRMPQAGVWSWEGCAGKRNKPVYTRLRPVNDMKEHTGV"

Query Match 100.0%; Score 3106; DB 10; Length 3106;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGTGAGACGGCTGCCCGAGGAGACCAAGCGCTGTAGAGACCGGCAATCGGACCGCCAAA 60
 DB 1 CATGTGAGACGGCTGCCCGAGGAGACCAAGCGCTGTAGAGACCGGCAATCGGACCGCCAAA 60
 QY 61 CCAATGGGTGACATTCGGGGCCGCAAGCCGAGGGGGCTCTCAGACTTCGGCGCGGAC 120
 DB 61 CCAATGGGTGACATTCGGGGCCGCAAGCCGAGGGGGCTCTCAGACTTCGGCGCGGAC 120
 QY 121 TCAAGTACAACTCCCGCTAGAGACATGATGGCTTTGAGAGGGTGTGAGATTCTTGC 180
 DB 121 TCAAGTACAACTCCCGCTAGAGACATGATGGCTTTGAGAGGGTGTGAGATTCTTGC 180
 QY 181 CTGCGAACAATGCCAAGAAAGTGAAGAGCGAGCCCGACGCGCTGGGTGTGCTGTG 240
 DB 181 CTGCGAACAATGCCAAGAAAGTGAAGAGCGAGCCCGACGCGCTGGGTGTGCTGTG 240
 QY 241 CAGTGTCTTCACTTCT 300
 DB 241 CAGTGTCTTCACTTCT 300
 QY 301 ATTATCGGAATGTGCGGGTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAATGAGA 360
 DB 301 ATTATCGGAATGTGCGGGTCAAAAAGTCTTCAATGGCCATCTGAGGATCACAATGAGA 360
 QY 361 TCTTTCTGATGGGTATGAGAACTCCACTCCACAGATTATTCAGCTTGGCCGACG 420
 DB 361 TCTTTCTGATGGGTATGAGAACTCCACTCCACAGATTATTCAGCTTGGCCGACG 420
 QY 421 TGAAGAGGGGGCTGAGAGCTGTGACATGAGTCCCTGCTGGGTCTCTACCAAGA 480
 DB 421 TGAAGAGGGGGCTGAGAGCTGTGACATGAGTCCCTGCTGGGTCTCTACCAAGA 480
 QY 481 AGTGGCTGTAACTGCTTCAAGTGAAGGAGTGTATGCGCTACTACTGTGTGAGATTCA 540
 DB 481 AGTGGCTGTAACTGCTTCAAGTGAAGGAGTGTATGCGCTACTACTGTGTGAGATTCA 540
 QY 541 GCATCCCCCACAAGCTGACAGAGAGGTGATGCGCCATGCGCTGAGAGGAGTGTGA 600
 DB 541 GCATCCCCCACAAGCTGACAGAGAGGTGATGCGCCATGCGCTGAGAGGAGTGTGA 600
 QY 601 CATTGCCACCCCGAGACCGGGCACTGAAATCTTGTGCTAACAATCTGTGTGCTTCC 660
 DB 601 CATTGCCACCCCGAGACCGGGCACTGAAATCTTGTGCTAACAATCTGTGTGCTTCC 660
 QY 661 CCATTGACCCAGATGTGTGACAGAGACTCAGACACAGCTGCAAGTTTGCCTGCATG 720
 DB 661 CCATTGACCCAGATGTGTGACAGAGACTCAGACACAGCTGCAAGTTTGCCTGCATG 720
 QY 721 CCCATGTGTGACAGTGTGACAGCTTCACTAACCCCTGGCTTCCCAACAGTCCCTAAC 780
 DB 721 CCCATGTGTGACAGTGTGACAGCTTCACTAACCCCTGGCTTCCCAACAGTCCCTAAC 780
 QY 781 CGCATGCCCCGAGTGGGTCTGCGGGGGGAGCGCGACTCTGTGTGAGCCCTCACT 840
 DB 781 CGCATGCCCCGAGTGGGTCTGCGGGGGGAGCGCGACTCTGTGTGAGCCCTCACT 840
 QY 841 TCCGAGCTTTGATGTGCTTCCCTGTGATGAGCATGGACGTGACTGTGTACCGTGTATG 900
 DB 841 TCCGAGCTTTGATGTGCTTCCCTGTGATGAGCATGGACGTGACTGTGTACCGTGTATG 900
 QY 901 ATAGCTTGAAGCCCATGAGACCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 DB 901 ATAGCTTGAAGCCCATGAGACCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 QY 961 CCTACAACTGACTTCTCTCTCTCCAGAAAGTCTTCTTGTGTACGCTGATTAACAAATA 1020

DB 961 CCTACAACTGACTTCTCTCTCTCCAGAAAGTCTTCTTGTGTACGCTGATTAACAAATA 1020
 QY 1021 CTGACCGGAGCATCTGCTTTGAGGCGACCTTCTTCAGCTGCGCCAGATGAGCACT 1080
 DB 1021 CTGACCGGAGCATCTGCTTTGAGGCGACCTTCTTCAGCTGCGCCAGATGAGCACT 1080
 QY 1081 GTGGCGGCTTTTGAAGTGAACCCCAAGGACATTTAGACGCGCTTACTATCCAGGCACT 1140
 DB 1081 GTGGCGGCTTTTGAAGTGAACCCCAAGGACATTTAGACGCGCTTACTATCCAGGCACT 1140
 QY 1141 ACCCGCCCAATCACTGACATGGAATATCAAGGTGCCCAACACCGGAGGTGAAG 1200
 DB 1141 ACCCGCCCAATCACTGACATGGAATATCAAGGTGCCCAACACCGGAGGTGAAG 1200
 QY 1201 TGGCTTCAAACTCTTCTATCTGTGTGAGACCCCAAGTACAGTGGCTCTCTGACCAAG 1260
 DB 1201 TGGCTTCAAACTCTTCTATCTGTGTGAGACCCCAAGTACAGTGGCTCTCTGACCAAG 1260
 QY 1261 ACTATGTGAGATCAACGGGAGAGATCTGCGGTGAGAGGTCCAGTTGTGTGAGCA 1320
 DB 1261 ACTATGTGAGATCAACGGGAGAGATCTGCGGTGAGAGGTCCAGTTGTGTGAGCA 1320
 QY 1321 GCAACGACGCAAGATTAAGTCCACTTCATCTGTATCACTGTATCAAGGACACGGGT 1380
 DB 1321 GCAACGACGCAAGATTAAGTCCACTTCATCTGTATCACTGTATCAAGGACACGGGT 1380
 QY 1381 TCTTAGGTGATGACTCTCTCTTCAAGCACTCCACAGACCGGTCCAGGATGTTATGTGA 1440
 DB 1381 TCTTAGGTGATGACTCTCTCTTCAAGCACTCCACAGACCGGTCCAGGATGTTATGTGA 1440
 QY 1441 AGACTGAGCGGTGATCCGAGAGAACTGCGTGCAGCGGCTGGGACAGCTGCCCGATT 1500
 DB 1441 AGACTGAGCGGTGATCCGAGAGAACTGCGTGCAGCGGCTGGGACAGCTGCCCGATT 1500
 QY 1501 ATAGTGTAGCGGTATCTGCGCATGGAATGCCACCCACAGTTTCAAGTGCAGAAAAACAGT 1560
 DB 1501 ATAGTGTAGCGGTATCTGCGCATGGAATGCCACCCACAGTTTCAAGTGCAGAAAAACAGT 1560
 QY 1561 TCTGCAAGCCCTCTTCTGCGGTGTGTGACAGTGTCAACGACTGTGGGACGGAAGTACG 1620
 DB 1561 TCTGCAAGCCCTCTTCTGCGGTGTGTGACAGTGTCAACGACTGTGGGACGGAAGTACG 1620
 QY 1621 AGAGGGCTGAGCTGTCTGTGAGAGATTCAAGTGTTCATATGGAAGTGTCTCTC 1680
 DB 1621 AGAGGGCTGAGCTGTCTGTGAGAGATTCAAGTGTTCATATGGAAGTGTCTCTC 1680
 QY 1681 AGAGCCGAAGTGTAAATGGGAAGGACAACTGTGAGATGGTCTGACGAGGCTTCACTGT 1740
 DB 1681 AGAGCCGAAGTGTAAATGGGAAGGACAACTGTGAGATGGTCTGACGAGGCTTCACTGT 1740
 QY 1741 ACAGCTGGAATGTGTCTCTTGCACCAATATTAACCTACCGCTGCACAAATGAGCTTGTG 1800
 DB 1741 ACAGCTGGAATGTGTCTCTTGCACCAATATTAACCTACCGCTGCACAAATGAGCTTGTG 1800
 QY 1801 TGAACAGAGGCAACCTGAGGTGTGATGGAGAGCGGACTGTAGCATGTGCTCGATGAGA 1860
 DB 1801 TGAACAGAGGCAACCTGAGGTGTGATGGAGAGCGGACTGTAGCATGTGCTCGATGAGA 1860
 QY 1861 AAAACTGTGATGTGGGTGTGAGATCTTACCAACAGGCTCGCTGTGTGTGTGTGTGTGT 1920
 DB 1861 AAAACTGTGATGTGGGTGTGAGATCTTACCAACAGGCTCGCTGTGTGTGTGTGTGTGT 1920
 QY 1921 ATGCGGACGAGAGGCGAGGTGGCGCTGTGAGAGGCTCCAGCGCTGAGGACAGGCGCACT 1980
 DB 1921 ATGCGGACGAGAGGCGAGGTGGCGCTGTGAGAGGCTCCAGCGCTGAGGACAGGCGCACT 1980
 QY 1981 TGTGTGGGCTCTGCTCATCTCTCTCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
 DB 1981 TGTGTGGGCTCTGCTCATCTCTCTCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2040
 QY 2041 ATGACAAATTTCAAGTACTGAGACTCAAGATGTGAGCGGCTTCTGGGTGTGTGTGTGT 2100

Db 2041 ATGACAAAATTTCAAGTACTCAGACTACACAGATGTGGAGCGGCTTCTGCTGTG 2100
 QY 2101 ACCAGAGCAGCCAGAGGCTCTGGGGTGCAGAGGCTGAAGCTCAAAAGTATCATCA 2160
 Db 2101 ACCAGAGCAGCGCAGGCTCTGGGGTGCAGAGGCTGAAGCTCAAAAGTATCATCA 2160
 QY 2161 ACCCTTCCCTTCAATGATTTACCTTGGCTGATATGATGCTGGAGCTGGAAGT 2220
 Db 2161 ACCCTTCCCTTCAATGATTTACCTTGGCTGATATGATGCTGGAGCTGGAAGT 2220
 QY 2221 CGGTGAGTACAGACCGTGTGGGCGCCCATCTGCTGCTGATGCTGATGCTGCTG 2280
 Db 2221 CGGTGAGTACAGACCGTGTGGGCGCCCATCTGCTGCTGATGCTGATGCTGCTG 2280
 QY 2281 CTGCTGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
 Db 2281 CTGCTGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
 QY 2341 CGCTGATCTGCGAGAGGGGTGAGATCGGTGCTATCAACAGACCACTGCTGAGACCT 2400
 Db 2341 CGCTGATCTGCGAGAGGGGTGAGATCGGTGCTATCAACAGACCACTGCTGAGACCT 2400
 QY 2401 TGCCTGAGCAGATCAACCCAGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 Db 2401 TGCCTGAGCAGATCAACCCAGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 QY 2461 CTGCTGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
 Db 2461 CTGCTGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
 QY 2521 AGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
 Db 2521 AGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
 QY 2581 CAAGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
 Db 2581 CAAGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
 QY 2641 AGACGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2700
 Db 2641 AGACGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2700
 QY 2701 GAACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2760
 Db 2701 GAACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2760
 QY 2761 GCATCCTTGAAGCTCAGAGTCTTCCAAAGTGGAGCAGCAGCAGCAGCAGCAGCAG 2820
 Db 2761 GCATCCTTGAAGCTCAGAGTCTTCCAAAGTGGAGCAGCAGCAGCAGCAGCAGCAG 2820
 QY 2821 TTGCTGCTAGCGGCGCCAGCTGCGGGGCAAGGCTTGAAGGCTTCCCTCTAGCC 2880
 Db 2821 TTGCTGCTAGCGGCGCCAGCTGCGGGGCAAGGCTTGAAGGCTTCCCTCTAGCC 2880
 QY 2881 CTGAGCTGAGTGAAGATGATGCTGCGGGAGGCTGCTCCAACTGCTGATGAGCTCC 2940
 Db 2881 CTGAGCTGAGTGAAGATGATGCTGCGGGAGGCTGCTCCAACTGCTGATGAGCTCC 2940
 QY 2941 GGGAGCCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3000
 Db 2941 GGGAGCCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3000
 QY 3001 ACCCGAGAGAGAGTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCT 3060
 Db 3001 ACCCGAGAGAGAGTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCTGAAGCT 3060
 QY 3061 TTGAGAGTAAACATTTATTTCTTTTAAAAA 3106
 Db 3061 TTGAGAGTAAACATTTATTTCTTTTAAAAA 3106

RESULT 3
 BC005496

LOCUS BC005496 3248 bp mRNA linear ROD 04-NOV-2003
 DEFINITION Mus musculus suppression of tumorigenicity 14 (colon carcinoma),
 mRNA (CDNA clone MGC:7395 IMAGE:3488059), complete cds.
 ACCESSION BC005496
 VERSION BC005496.1 GI:13529565
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3248)
 1 Klausner,R.D., Colling,L.E., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Zeeberg,B., Burow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stappleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,V., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
 Dickinson,M.C., Rodriguez,A.C., Grimwood,J., Schmitz,J., Myers,R.M.,
 Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalton,D.B.,
 Schermer,A., Schein,J.E., Jones,S.J., and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 3248)
 Strausberg,R.
 Direct Submission
 Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Leohar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAK Plate: 7 Row: 1 Column: 16
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 FEATURES
 source
 Location/Qualifiers
 1..3248
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="MGC:7395 IMAGE:3488059"
 /tissue_type="mammary tumor, WAP-TGF alpha model, 7 months
 old, gross tissue."
 /clone_id="NCI CGAP_Mam5"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..3248
 gene

QY 1321 GCAACAGCAGCAAGATTACAGTCCATTCCATTCTGATCACTGCTACACGACACCGGGT 1380
 Db 1431 GCAACAGCAGCAAGATTACAGTCCATTCCATTCTGATCACTGCTACACGACACCGGGT 1490
 QY 1381 TCTTAGTGAATCCTCTCTCTCAAGCTCAACGACCCGGTCCAGGAGATGTTATGTCGA 1440
 Db 1491 TCTTAGTGAATCCTCTCTCTCAAGCTCAACGACCCGGTCCAGGAGATGTTATGTCGA 1550
 QY 1441 AGACTGGAAGCTGATCCGAAAGAACTGCTGCGACGAGCTGCGGAGACTGCGCGAGT 1500
 Db 1551 AGACTGGAAGCTGATCCGAAAGAACTGCTGCGACGAGCTGCGGAGACTGCGCGAGT 1610
 QY 1501 ATAGTATGAGCGTTACTGCGGATGCAATGCGACCCACCGATTCACTGCAAAAACCACT 1560
 Db 1611 ATAGTATGAGCGTTACTGCGGATGCAATGCGACCCACCGATTCACTGCAAAAACCACT 1670
 QY 1561 TCTGCAAGCCCTCTCTCTGAGTCTGACAGTCTCAACGACTGCGGAGACGAAAGTGAAG 1620
 Db 1671 TCTGCAAGCCCTCTCTCTGAGTCTGACAGTCTCAACGACTGCGGAGACGAAAGTGAAG 1730
 QY 1621 AGGAGGGCTGCACTGCTCTCTGAGGATTTCAAGTTCGAAAGGAAAGTGTCTCTCTC 1680
 Db 1731 AGGAGGGCTGCACTGCTCTCTGAGGATTTCAAGTTCGAAAGTGTGTCTCTCTC 1790
 QY 1681 AGAGCCAGAGTGTATGAGAAAGACAACTGTGAGATGAGTGTGACGAGGCTTCAATG 1740
 Db 1791 AGAGCCAGAGTGTATGAGAAAGACAACTGTGAGATGAGTGTGACGAGGCTTCAATG 1850
 QY 1741 ACAGCGTGAATGTCTCTCTGACCAAAATATCTACCGCTGCGCAAAATGAGCTGTCTC 1800
 Db 1851 ACAGCGTGAATGTCTCTCTGACCAAAATATCTACCGCTGCGCAAAATGAGCTGTCTC 1910
 QY 1801 TGAAGCAAGGCGCACTCTGATGATGAGAAAGCGAGCTGTAGCGATGCTCCGATGAGA 1860
 Db 1911 TGAAGCAAGGCGCACTCTGATGATGAGAAAGCGAGCTGTAGCGATGCTCCGATGAGA 1970
 QY 1861 AAAACGTGATCTGTGGGCTGCGATCTTTACCAACAGGCTGCGGTGTGTGGACGA 1920
 Db 1971 AAAACGTGATCTGTGGGCTGCGATCTTTACCAACAGGCTGCGGTGTGTGGACGA 2030
 QY 1921 ATGCGGACGAGGCGAGTGGCCCTGAGTGGAGCTTCCACGCGCTGCGGCGCAGGCGACT 1980
 Db 2031 ATGCGGACGAGGCGAGTGGCCCTGAGTGGAGCTTCCACGCGCTGCGGCGCAGGCGACT 2090
 QY 1981 TGTGTGGGCGCTGCTCATCTCTCTGACTGCTGCTGTCTGTGACGCTATGCTTTACG 2040
 Db 2091 TGTGTGGGCGCTGCTCATCTCTCTGACTGCTGCTGTCTGTGACGCTATGCTTTACG 2150
 QY 2041 ATGACAAAATTTCAGTACTGAGACTACAGATGAGAGCGGCTTCTGCGGTCTGCTGG 2100
 Db 2151 ATGACAAAATTTCAGTACTGAGACTACAGATGAGAGCGGCTTCTGCGGTCTGCTGG 2210
 QY 2101 ACCAGAGCAAGGCGAGTCTCTGAGGAGTGAAGCTGAAAGCTGAAAGCTATCATCACCC 2160
 Db 2211 ACCAGAGCAAGGCGAGTCTCTGAGGAGTGAAGCTGAAAGCTGAAAGCTATCATCACCC 2270
 QY 2161 ACCCTTCTTCAATGATTTCACTTCACTATGACATGCTGCTGTGAGCTGAGAGAGT 2220
 Db 2271 ACCCTTCTTCAATGATTTCACTTCACTATGACATGCTGCTGTGAGCTGAGAGAGT 2330
 QY 2221 CGGTGAGTACAGACCGTCTGCGGCGCCCATCTGCTGCTGAGTGTGCTTCTTCC 2280
 Db 2331 CGGTGAGTACAGACCGTCTGCGGCGCCCATCTGCTGCTGAGTGTGCTTCTTCC 2390
 QY 2281 CTGCTGAGCAAGCCATCTGGGTCAAGGCTGAGGCGCACAAAAGAGGAGGTACCGGAG 2340
 Db 2391 CTGCTGAGCAAGCCATCTGGGTCAAGGCTGAGGCGCACAAAAGAGGAGGTACCGGAG 2450
 QY 2341 CGCTGATCTTCAAGAGGCTGATCGTGTCTATCAACAGACCACTGTGAGGACCTCA 2400
 Db 2451 CGCTGATCTTCAAGAGGCTGATCGTGTCTATCAACAGACCACTGTGAGGACCTCA 2510
 QY 2401 TGCCGAGCAGATACCCCAAGATGATGTGTGTGGGTTCTCTAGTGGGCTGTGAGCT 2460

Db 2511 TGCCGAGCAGATACCCCAAGATGATGTGTGGTTTCTCAAGTGGGCTGTGAGCT 2570
 QY 2461 CTGCGCAGGCTGATCTGTGTGGCCCTTGTGCAAGCGCGAGAAAGATGCGATGTTCC 2520
 Db 2571 CTGCGCAGGCTGATCTGTGTGGCCCTTGTGCAAGCGCGAGAAAGATGCGATGTTCC 2630
 QY 2521 AGGCTGATGTGAGCTGCGGAGTGAAGGCTGCGCTGAGAGGAAACAGAGGCTGTACA 2580
 Db 2631 AGGCTGATGTGAGCTGCGGAGTGAAGGCTGCGCTGAGAGGAAACAGAGGCTGTACA 2690
 QY 2581 CAAAGCTCCTGTAGTTCGAGACTGATCAAAAGACACTGCGGTATATACAGATGAGAC 2640
 Db 2691 CAAAGCTCCTGTAGTTCGAGACTGATCAAAAGACACTGCGGTATATACAGATGAGAC 2750
 QY 2641 AGACACCCAGCCCAAAACCCACAGAGATGCCGACATGACACCTGTATACAGAGAG 2700
 Db 2751 AGACACCCAGCCCAAAACCCACAGAGATGCCGACATGACACCTGTATACAGAGAG 2810
 QY 2701 GAACACTGACGACATTTATGCTGTGGCTCCCGCCCGCAACACACCGAGCTGTGACT 2760
 Db 2811 GAACACTGACGACATTTATGCTGTGGCTCCCGCCCGCAACACACCGAGCTGTGACT 2870
 QY 2761 GCATCCTTTAGACTCAGAGTTCCTTCCAAAGTGGACCCCTCAAGAGTTGAGAGAGAAC 2820
 Db 2871 GCATCCTTTAGACTCAGAGTTCCTTCCAAAGTGGACCCCTCAAGAGTTGAGAGAGAAC 2930
 QY 2821 TTGCGGTGACAGGCGCCACGCTGCGGCGGACAGGCTTTGATGAGACCTTCCCTTAC 2880
 Db 2931 TTGCGGTGACAGGCGCCACGCTGCGGCGGACAGGCTTTGATGAGACCTTCCCTTAC 2990
 QY 2881 CTGAGCTGGGTGAGATGATGTGTCTCCGAGAGAGCTGCTTCAACTGTATGAGTCCC 2940
 Db 2991 CTGAGCTGGGTGAGATGATGTGTCTCCGAGAGAGCTGCTTCAACTGTATGAGTCCC 3050
 QY 2941 GGGAGCCCTTATGAGAGAGGCGCTCAGGCTCACTCTTTTACAGAAAGCCGACCTTACGA 3000
 Db 3051 GGGAGCCCTTATGAGAGAGGCGCTCAGGCTCACTCTTTTACAGAAAGCCGACCTTACGA 3110
 QY 3001 ACCCGAGAAAGAGTGTATACCTTAAAGCTGAAATTTGTTTCTGTTGCCAGGCTGTAT 3060
 Db 3111 ACCCGAGAAAGAGTGTATACCTTAAAGCTGAAATTTGTTTCTGTTGCCAGGCTGTAT 3170
 QY 3061 TTGAGAGTAAACATTTATTTCTTTTAAAAAAA 3106
 Db 3171 TTGAGAGTAAACATTTATTTCTTTTAAAAAAA 3216

RESULT 4
 AB049189 3174 bp mRNA linear ROD 14-APR-2001
 LOCUS
 DEFINITION Rattus norvegicus mRNA for membrane bound arginine specific serine
 protease, complete cds.
 AB049189
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (sites)
 2 (bases 1 to 3174)
 TITLE
 JOURNAL
 INoue, H., Takahashi, K. and Kishi, K.
 membrane-bound arginine specific serine protease
 Published Only in Database (2000)
 INoue, H., Takahashi, K. and Kishi, K.
 Direct Submission
 Submitted (22-SEP-2000) Hideaki Inoue, Tokyo University of Pharmacy
 and Life Science, School of Life Science, 1432-1 Horiinouchi,
 Hachioji-shi, Tokyo 192-0392, Japan (E-mail: hinoue@is.toyaku.ac.jp,
 Tel: 81-426-76-7153, Fax: 81-426-76-7157)
 Location/Qualifiers


```

source
1. .3174
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="duodenum"
/dev_stage="7 weeks old"
127..2694
/codon_start=1
/evidence=experimental
/product="membrane bound arginine specific serine
protease"
/protein_id="BAB13765.1"
/db_xref="GI:10336527"
/translation="MGNNRGRKAGGGSDFGAGLKYNSLNNNGEBSVEVELPVNNA
KQVEKGRPRVWVAVVPSFLILSLMAGLWHPHYRVRLOKXNGHLRTNEFL
DAVENSTSEFISLQVKEALMYSEVPLGPYHKSTVLAFSGSIYAWSEFS
IPPHLEERYDRMAVRVVTLPARALSPVLSVAVPDPIDRMLORTDNCSFAL
HARGTVTRFTTPGFPNSYPARARCOMVTLRGDSDVLSLTRSPVARCDGDSLV
TVYDLSLPEPHAVRVLGTFSESTNLTLSSQNVFLVLTITDRHRHGFATPRL
PVMSGCGALISEQTFSSPYPGHYPEINICTWNLKPVNRNVKRFKFLVLPDNI
PVGCTKDYVLEINKEKFCGERSQFVVSNSKLTVAHSDHSTYDGLAEYISYDNI
DPCPEMCKTGRCIRKLRCQDMDPCPYSDERHCRNATQPMCKNOFCRFLFVVC
DSVNDGDSDEBSCSPAGSEKSCKLPDQOCNGDKDGLRSTKQARVVGTMADG
CTKYTRCONGLINKNPPCDGKDCDSCDEKNCDCGLRSTKQARVVGTMADG
EPMQVSLHAGGHLGASLISPDMLVSAHCFQDETYFKYSDHMTMTAFGLDLS
KRSASGVQEHKRLRIITPSEFNDFTYDIALLEKRPAYSTVYRPLCPDTHYFP
AGKAIWVTGHTKEGTGALILQKGIIRINQTEELLPQDITRRMNCVGLSGV
DSCGDSGGLSVEKGRIFQAGVVSWEBCAQNRNKPVTYRIPEVRMIKEQTV"
360
variation
/note="compared with the sequence reported in Acc#
AB037898"
/replace="t"
513
variation
/note="compared with the sequence reported in Acc#
AB037898"
/replace="t"
573
variation
/note="compared with the sequence reported in Acc#
AB037898"
/replace="t"
2121
variation
/note="compared with the sequence reported in Acc#
AB037898"
/replace="t"
3157. .3174
polysite
ORIGIN
Query Match 83.5%; Score 2595; DB 10; Length 3174;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 2853; Conservative 0; Mismatches 230; Indels 50; Gaps 7;

```

```

QY 301 ATTATGGAATGTGCGGTTCAAAAAAGTCTTCATGGCCATCTGAGGATCACAAAATGAGA 360
DB 365 ATTACGGGAATGTAGGATCCAAAAAGTCTTCATGGCCATCTGAGGATCACAAAATGAGA 424
QY 361 TCTTTCTGGATCGGATGAGAACTCACTCCACAGAGTTTATCGAGCTGGCGAGCAGG 420
DB 425 ACTTTCTGACCGGATGAGAACTCACTCCACAGAGTTTATCGAGCTGGCGAGCAGG 484
QY 421 TGAAGAGGCGTGAAGCTGTGTAATGAAGTCCCTGCTGGTCCCTACCAAGA 480
DB 485 TGAAGAGGCGTGAAGCTGTGTAATGAAGTCCCTGCTGGTCCCTACCAAGA 544
QY 481 AGTCGGCTGTAATCTGCTTCAGTGAAGGCACTGTATCGCTTACTCTGTGAGAGTTCA 540
DB 545 AATCGACTTAATCTGCTTCAGTGAAGGCACTGTATCGCTTACTCTGTGAGAGTTCA 604
QY 541 GCATCCCCCAACCTGGGAGAGAGTTGATCGGCGCATGGCTGGAGGCGAGTTGTA 600
DB 605 GCATCCCCCAACCTGGGAGAGAGTTGATCGGCGCATGGCTGGAGGCGAGTTGTA 664
QY 601 CATTCGCACCCCGAGACCGGGCACTGAAATCCTTGTGCTTAACATCTGTGGGCTTCC 660
DB 665 CACTGCAACCCGAGACCGGGCACTGAAATCCTTGTGCTTAACATCTGTGGGCTTCC 724
QY 661 CCATTGACCCGAGATGCTGAGAGGACTCAGAGACAACAGTGCAGTTTGGCTGATG 720
DB 725 CCATTGACCCGAGATGCTCAGAGGACTCAGAGACAACAGTGCAGTTTGGCTGATG 784
QY 721 CCATAGTGAGAGAGAGACAGCTTCACTACCTCCCTGCTTCCCAACAGTCCCTACCCG 780
DB 785 CCGTGACAGAGACAGAGACAGTTTCACTACCTCCCTGCTTCCCAACAGTCCCTACCCG 844
QY 781 CGCATGCCGCTGCGAGTGGTCTGTGGGGGAGACGCGGACTGTGCTGAGCTTCACT 840
DB 845 CTGATGCCGCTGCGAGTGGTCTGTGGGGGAGACGCGGACTGTGCTGAGCTTCACT 904
QY 841 TCCGAAGCTTGAATGTGCTGCTTCTGTGATGACATGAGACCTGTGATGACCTGTATG 900
DB 905 TCCGAAGCTTGAATGTGCTGCTTCTGTGATGAGATGAGACCTGTGATGACCTGTATG 964
QY 901 ATGAGCTGAGCCCATGAGACCCGACGCTGTGGGCGGCTGTGAGCACTTCTCACT 960
DB 965 ACAGCTGAGCCCATGAGACCCGACGCTGTGGGCGGCTGTGAGCACTTCTCACT 1024
QY 961 CTTACAACTGACTTCTCTCTCTCCAGAAAGCTTCTTTGTCAAGCTGATTAACAATA 1020
DB 1025 CTTACAACTGACTTCTCTCTCTCCAGAAAGCTTCTTTGTCAAGCTGATTAACAATA 1084
QY 1021 CTGACCGGCGACATCTGCTTGTGAGGCACTTCTTCCAGCTGCGCAAGATGAGAGCT 1080
DB 1085 CTGACCGGCGACATCTGCTTGTGAGGCACTTCTTCCAGCTGCGCAAGATGAGAGCT 1144
QY 1081 GTGGCGGCTTTTGAAGGACACCCAGAGGACATTTAGAGGCTTACTATCCAGGCACT 1140
DB 1145 GTGGCGGCTTTTGAAGGACACCCAGAGGAGCTTTAGAGGCTTACTATCCAGGCACT 1204
QY 1141 ACCGCGCCCAATCACTGCACTGATGAATATCAAGGTGCCCCAACAACCGGAAGTGAAG 1200
DB 1205 ACCGCGCCCAATCACTGCACTGATGAATATCAAGGTGCCCCAACAACCGGAAGTGAAG 1264
QY 1201 TGGCTTCAAACTTCTCTATCTGTGAGACCCCAAGTACAGTGGGCTCTGCAACAAG 1260
DB 1265 TACGCTTCAAACTTCTCTATCTGTGAGACCCCAAGTACAGTGGGCTCTGCAACAAG 1324
QY 1261 ACTATGTGAGATCAACGGGAGAGATCTCGGTGAGAGGTCCAGTTTGTGTGACA 1320
DB 1325 ACTATGTGAGATCAACGGGAGAGATCTCGGTGAGAGGTCCAGTTTGTGTGACA 1384
QY 1321 GCACAGCAGCAAGATTACAGTCACTTTCATCTGATCACTGTACACGACACCGGGT 1380
DB 1385 GCACAGCAGCAAGATTACAGTCACTTTCATCTGATCACTGTACACGACACCGGGT 1444
QY 1381 TCTTAGCTGAGTACCTCTCTTACGACTCCACGACCGGTCGCCAAGAGATTCAATGTGA 1440

```


Laboratory of Nutrition Chemistry, Oiwake-cho, Kitashirakawa,
Sakyo-ku, Kyoto 606-8502, Japan
(E-mail: tkhono@kais.kais.kyoto-u.ac.jp, Tel: 81-75-753-6263,
Fax: 81-75-753-6264)

FEATURES
source
Location/Qualifiers

1. 2568
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="jejunum"
/dev_stage="8 weeks old"
1. 2568
/gene="MBSP"
1. 2568
/gene="MBSP"
/codon_start=1
/product="membrane bound serine protease"
/protein_id="BAB03502.1"
/db_xref="GI:9650964"
/translation="MGNNRGRKAGGGSDPAGLKYNSRLNNMGEEVEPELPVNA
/KVEKGPRLVAVVPSFLISIMAGLVHFRVRLQKVENGLRTNENFL
DAENSTSEFISLQVKALIMYSEVLPYHKSLVAFSGSVIAYWSEFS
IPHLEEDRAMAVRVTLPARALSFVLTSVAFIPDRMDORTDNCSFAL
HARGVTRPTTTPGPNSEYPARARCOMVIRGADSVLSLTPRFPVAPDGDSDLV
TVYDLSLPEHNAVRLCGTFSPSYNLTFLSSQNVFLVLTITPDRHREPEATPOL
PKMSGCGGLISEAGTFSSPYPGHYHPNINCTWNIKVPNNRVAKRFKFLVYDNI
PVSGCTDYIEINGEFCGERSQVVSNSKLTFFHSDHSYTDGFLAELYSYDN
DPCKMCKTGKICRKLDCWADCPVDSRHRCAVTHDPMCKNPFCKLFWCS
DYNDCDSDERGCSPAGSFCKNSGKCLPOSQCNKGDGSDGDEASCDNVAVS
CTRYTRCQNGCLTANKNPECDKDCSDSDERKNDGCRSTFKARVAGTNADEG
EMWQVSLH/LGQHLGASLISPDMLVSAHCFQDPTIKYSDHMTATFLILLOS
KRASGVQEHKRLIITHPSEFNDFTPYDIALLEKPAVSTVAPRICLPDTHNP
AGAAIWTGHTREGTGALILQKGRIVINQITCEELPQOITPMMCVGLSGV
DSCQDSGGLSVEKGRIFQAGVWSMBCAQRNRPVYTRIIPVDRWIKQTV"

ORIGIN

Query Match 74.5%; Score 2313.6; DB 10; Length 2568;

Best Local Similarity 93.8%; Pred. No. 0; Mismatches 159; Indels 0; Gaps 0;

Matches 2409; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

63 ATGGGTGCAATCGGGCCGCAAGCGGAGGGGCTCTCAGACTTCGGCGCGGACTC 122
1 ATGGGAAACAATCGGGCCGCAAGCGGAGGGGCTCTCAGACTTCGGCGCGGACTC 60
123 AAGTACAATCCCGGCTAGAGACATGATGCTTTGAGAGGGGTGTGAGATTCTGCT 182
61 AAGTACAATCCCGGCTAGAGACATGATGCTTTGAGAGGGGTGTGAGATTCTGCT 120
183 GCGAACAATGCCAAGAGTGAAGAGGAGCCCGAGCGCTGGGTGTGAGTGTGCA 242
121 GTGAACAACGCAAGCAGTGAAGAGGAGCCCGAGCGCTGGGTGTGATGATGATG 180
243 GTGCTGTGAGCTTCTCTGCTCTCTCCTCATGCTGCTGCTGCTGCTGCTGCTCAT 302
181 GTAGTGTGAGCTTCTCTGCTCTCTCCTCATGCTGCTGCTGCTGCTGCTGCTCAT 240
303 TATCGAATGTGCGGGTCAAAAAGTCTTCAATGAGCCATCTGAGATCAAAATGATC 362
241 TACGGAAATGTGAGGATCCAAAAGTCTTCAATGAGCCATCTGAGATCAAAATGATC 300
363 TTTCTGATGGGTATGAGAACTCCCACTCCCAAGTTTATCAGCTGGCCAGCGAGTG 422
301 TTTCTGAGCGGTATGAGAACTCCCACTCCCAAGTTTATCAGCTGGCCAGCGAGTG 360
423 AAGGAGCGCTGAGCTGCTGATCAATGAGCTCTGCTGAGCTCTCTTACCAAGAAG 482
361 AAGGAGCGCTGAGCTGATGAGTGAAGTCCCTGCTGAGCGCTCTTACCAAGAAG 420
483 TCGGCTGATGCTCTTCAAGTGAAGGCAAGTGTGATGCTGCTTACTGCTGAGATTGAC 542
421 TCGAATGATGCTCTTCAAGTGAAGGCAAGTGTGATGCTGCTTACTGCTGAGATTGAC 480

543 ATCCCCCAGACCTGAGCAAGAGGTTGATCGCGCCATGCTGTGAGGCAAGTTGACA 602
481 ATCCCCCAGACCTGAGCAAGAGGTTGATCGCGCCATGCTGTGAGGCAAGTTGACA 540
603 TTGCAACCCCGAGCAAGGCACTGAATCTCTGCTGCTCAATCTGTGAGGCTTCCGC 662
541 CTGCAACCCCGAGCAAGGCACTGAATCTCTGCTGCTCAATCTGTGAGGCTTCCGC 600
663 ATTGACCCCGAGATGCTGAGAGGACTGAGCAACAGAGCTGACAGTTTCCCTGATGAC 722
601 ATTGACCCCGAGATGCTGAGAGGACTGAGCAACAGAGCTGACAGTTTCCCTGATGAC 660
723 CATGTGACAGAGTGAACGCTTCACTAACCCCTGCTTCCCAACAGTCCCTAACCCGCG 782
661 CATTGAGAGAGAGTGAACGCTTCACTAACCCCTGCTTCCCAACAGTCCCTAACCCG 720
783 CATGCCGCTGCGAGTGGTCTGCGCGGGAGAGCCGCACTCTGCTGAGGCTCACCTTC 842
721 CATGCCGCTGCGAGTGGTCTGCGCGGGAGAGCCGCACTCTGCTGAGGCTCACCTTC 780
843 CGAAGCTTGAATGCTGCTGCTGCTGATGAGCATGAGTACCTGATCACTGATGAT 902
781 CGAAGCTTGAATGCTGCTGCTGCTGATGAGCATGAGTACCTGATCACTGATGAT 840
903 AGCCTGAGCCCAATGGAACCCCAAGCTGTGAGGCTGCTGAGCACTTCTCACTCC 962
841 AGCCTGAGCCCAATGGAACCCCAAGCTGTGAGGCTGCTGAGCACTTCTCACTCC 900
963 TACAACCTGATCTTCTCTCTCCCAAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCT 1022
901 TACAACCTGATCTTCTCTCTCCCAAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCT 960
1023 GACCGGAGACATCTGCTGCTTGAAGCCACTTCTTCCAGCTGCGCAAGATGAGAGCTGT 1082
961 GACCGGAGACATCTGCTGCTTGAAGCCACTTCTTCCAGCTGCGCAAGATGAGAGCTGT 1020
1083 GAGCGCTTTTGAATGACACCCCAAGGACATTTAGAGCCCTTACTATCCAGGCTACTAC 1142
1021 GAGCGCTTTTGAATGACACCCCAAGGACATTTAGAGCCCTTACTATCCAGGCTACTAC 1080
1143 CCGCCCAACATCAATGCAATGGAATATGAGAGTGGCCCAACACCGGAACGTGAAGTA 1202
1081 CCGCCCAACATCAATGCAATGGAATATGAGAGTGGCCCAACACCGGAACGTGAAGTA 1140
1203 CGCTTCAACTCTTCTATCTGCTGAGCCCAACGCTACAGTGGCTCTTGCACCAAGAC 1262
1141 CGCTTCAACTCTTCTATCTGCTGAGCCCAACGCTACAGTGGCTCTTGCACCAAGAC 1200
1263 TATGTGAGATCAACGCGGAGAGATCTGCGGTGAGAGTCCCAAGTTTGTGTAGACAGC 1322
1201 TATGTGAGATCAACGCGGAGAGATCTGCGGTGAGAGTCCCAAGTTTGTGTAGACAGC 1260
1323 AAGCAGCAGAGATTAAGATCACTTCATCTGATCACTGTCACAGCAACCGGGTTTC 1382
1261 AAGCAGCAGAGATTAAGATCACTTCATCTGATCACTGTCACAGCAACCGGGTTTC 1320
1383 CTAGCTAGTACTCTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1442
1321 CTAGCTAGTACTCTCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1380
1443 ACTGAGCGGTGATCCGAAGAGAGTGGCTGCGAGCGCTTGGGCAAGATGCGCCGATAT 1502
1381 ACTGAGCGGTGATCCGAAGAGAGTGGCTGCGAGCGCTTGGGCAAGATGCGCCGATAT 1440
1503 AGTATGAGCGGTGATCCGAAGAGAGTGGCTGCGAGCGCTTGGGCAAGATGCGCCGATAT 1562
1441 AGTATGAGCGGTGATCCGAAGAGAGTGGCTGCGAGCGCTTGGGCAAGATGCGCCGATAT 1500
1563 TCGAAGCCCTCTTCTGAGTGTGAGCAAGTGTCAACGCTGTGAGGAGAGGAGTGAAG 1622
1501 TCGAAGCCCTCTTCTGAGTGTGAGCAAGTGTCAACGCTGTGAGGAGAGGAGTGAAG 1560

QY	1623	GAGGCGTCGACGCTGTCCTCTGGAGATTTCAAGTGTTCATGGAGAGTGTCTCCCTCAG	1682
Dp	1561	GAGGCGTCGACGTTGTCTCTCCGGAGTTTCAAGTGTTCATGGAGAGTGTCTCCCTCAG	1620
QY	1683	AGCCAGAACTGTATATGGGAAGACACTGTGGAGATGGGTCTGAGAGAGCTTCATGTGAC	1742
Dp	1621	AGCCAGCACTGTATATGGGAAGACACTGTGGAGATGGGTCTGAGAGAGCTTCATGTGAC	1680
QY	1743	AGCGTGAATGTGTCTCTTGGACCAATATATACCTACCGCTGCATAATATGAGCTCTGTCTG	1802
Dp	1681	AATGTGAATGCCGTCTCTGACCAAAATATATACCTACCGCTGCATAAATATGAGCTCTGTCTG	1740
QY	1803	AGCAAGGGCAACCCCTGAGTGTATATGGGAAGACGACTGTAGCATGCTCCGATGAGAAA	1862
Dp	1741	AACAAAGGGCAACCCCTGAGTGTATATGGGAAGACGACTGTAGCATGCTCCGATGAGAAA	1800
QY	1863	AACTGTGACTGTGGGCTGGATCCTTTACAAAAGGCTCGAGTGTGGTGGGACCAAT	1922
Dp	1801	AACTGTGACTGTGGGCTGGATCCTTTACAAAAGGCTCGAGTGTGGTGGGACCAAT	1860
QY	1923	GGGACAGAGGCGAGTGGCCCTGGCAGGTGAGCTTCAAGCCCTGGGCGAAGGCCACTTG	1982
Dp	1861	GGGACAGAGGCGAGTGGCCCTGGCAGGTGAGCTTCAAGCCCTGGGCGAAGGCCACTTG	1920
QY	1983	TGTGGGGGCTCGCTCATCTCTCCCTGACGCGTGTCTGTGAGCTCATGTGTTCAAGAT	2042
Dp	1921	TGTGGGGGCTCGCTCATCTCTCCCTGACGCGTGTCTGTGAGCTCATGTGTTCAAGAT	1980
QY	2043	GACAAAAATTTCAAGTACTACACTACAGATGTGACAGGCTTCTGGGTCTGTGGAC	2102
Dp	1981	GAGACAAATTTCAAGTACTACAGACCAACATGTGACAGCTTCTGGGTCTGTGGAC	2040
QY	2103	CAGACCAAGCCGAGGCTCTTGGGGTGTAGAGAGCTGAACTCAAACTATCATACCCAC	2162
Dp	2041	CAGACCAAGCCGAGGCTCTTGGGGTGTAGAGAGCAAGCTCAAACTATCATACCTAC	2100
QY	2163	CGTTCTCTTCAATGATTTACCTTGACATATACATCGCTGCGTGGAGCTGGAGAAAGTGG	2222
Dp	2101	CGTTCTCTTCAATGATTTACCTTGACATATGACTTGTGAGCTTGTGGAGAAAGCTGG	2160
QY	2223	GTGGAGTACAGACCGCTGTGTCGCCCCCATCTGCTGTATGTACTCAATGTCTTCCCT	2282
Dp	2161	GCAAGATACAGACCTGTGTGTCGCCCCCATCTGCTGTATATATCCATGTCTTTCCT	2220
QY	2283	GCTGGCAAGGCGCATCTGGGTACACAGGCTGGGGGCGACACAAABAAGGAGGTACCGGAGCG	2342
Dp	2221	GCCGGCAAGGCGCATCTGGGTACACAGGCTGGGGGCGACACGAAABAAGAGGAACTTGAGACA	2280
QY	2343	CTGATCTTGCAGAAAGGGTGAGATCCGTGATCAACACAGACCACTGTGAGAGCTTCATG	2402
Dp	2281	CTGATCTTGCAGAAAGGAGAGATCCGGGTATCAACAGACCACTGTGAGAACTTCCTG	2340
QY	2403	CCGAGAGAGATCAACCCCAAGATATGTGTGTGGGTTTCTCATGTGGGGGTGTGACTCC	2462
Dp	2341	CCGAGAGAGATCAACCCCAAGATATGTGTGTGGGTTTCTCATGTGGGGGTGTGACTCC	2400
QY	2463	TGCCAGGGTACTCTGTGTGCGCCCTTGTCAAGGCGGAGAAABAATGGGCGAAATGTTCCAG	2522
Dp	2401	TGCCAGGGTATATTCGGTGTGCGCCCTTGTCAAGTGTGAGAAABAATGGGCGAAATCTTCCAG	2460
QY	2523	GCTGTGTGTGTGACTGTGGGTGAAAGCTGTGCTCAAGAGAACAAAGCCAGCGTGTACACA	2582
Dp	2461	GCTGTGTGTGTGACTGTGGGTGAAAGCTGTGCTCAAGAGAACAAAGCCAGCGTGTACACA	2520
QY	2583	AGGCTCCCTGTAGTGTGGGACTGGATCAAGACACACTGGGGTATATG	2630
Dp	2521	AGGATCCCTTAAGTGTGGGACTGGATCAAGACCAACTGGGGTATATG	2568

FEATURES	source
ACCESSION	matpitase, epithin), mRNA (CDNA clone MGC:40392 IMAGE:52131389), complete cds.
VERSION	BC030532
KEYWORDS	BC030532.1 GI:20988874
SOURCE	MGC.
ORGANISM	Homo sapiens (human)
REFERENCE	Homio sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 3273) Strausberg,R.L., Peirngold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shemmen,C.M., Schuler,G.D., Altchul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stalcenon,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carrion,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Whlathy,S.J., Bosak,S.A., McMan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzyzanski,M.I., Skalska,U., Smalits,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PMID	22388257
REFERENCE	12477932
AUTHORS	2 (bases 1 to 3273) Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland. Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@hgrl.nih.gov Ahtler,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Hsieh,H.P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Loric,P., Legaspi,R., Madun,Q.L., Mastello,C., Maekel,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W., Tsurgan,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image-llnl.gov Series: IRAC Project: 64 Row: k Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein. Location/Qualifiers 1..3273

RESULT 6	LOCUS	DEFINITION
BC030532	3273 bp	linear
	RNA	PRI 07-OCT-2003
	suppression of tumorigenicity 14 (colon carcinoma,	

clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://image.lnl.gov>

Series: IRAK Plate: 64 Row: K Column: 15

This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

Location/Qualifiers

1..3273

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:40392 IMAGE:5213189"

/tissue_type="Blood, adult leukocytes"


```

Db      1316 ACTAGTGGAGATCAATGGGGAATAATCTGGAGAGAGGTCCAGTCTCTGTCAACA 1375
Qy      1321 GCAACAGCAGAGATTACAGTCCACTTCCATTCTGTATCACTGTAGACGAGACCGGGT 1380
Db      1376 GCAACAGCAGAGATTACAGTCCACTTCCATTCTGTATCACTGTAGACGAGACCGGGT 1435
Qy      1381 TCCAGAGTGAATCTCTCTGAGCTCCAGACGAGACCGGGTCCAGAGATTTGATGCA 1440
Db      1436 TCTTAGTGAATCTCTCTGAGCTCCAGACGAGACCGGGTCCAGAGATTTGATGCA 1495
Qy      1441 AGACTGAGCGATCCAGAAAGGAATGGCGTCCGAGCGGTGGGAGACTGCGCGGAT 1500
Db      1496 GCAACGGGCGGTATTCGAGAGAGAGCTGGCTGTGATGCTGGCGAGCTGACCGACC 1555
Qy      1501 ATAGTATGAGCGTTACTGCGATGCAATGCCACCCACAGTTCACTGTGCAAAACAGT 1560
Db      1556 ACAGCGATGAGCTCAACTGCAATGGGAGCGCGGCCACCGATTCACTGTGCAAAACAGT 1615
Qy      1561 TCTGCAAGCCCCCTTCTGTGGGTCTGTGACAGTGTCAAAGCTGTGGGAGCGAAGTACG 1620
Db      1616 TCTGCAAGCCCCCTTCTGTGGGTCTGTGACAGTGTCAAAGCTGTGGGAGCGAAGTACG 1675
Qy      1621 AGAGAGGCTGACAGTCTGCTGTGGAGATTCAAGTTCATAGGAAATGTCTCCCTC 1680
Db      1676 AGAGAGGCTGACAGTCTGCTGTGGAGATTCAAGTTCATAGGAAATGTCTCCCTC 1735
Qy      1681 AGAGCGAAGTGTAAATGGGAGAGCACTGTGAGATGGATGTCAGAGGCTTCAATGTG 1740
Db      1736 AAAGCCAGAGTGCATGGGAAAGAGCAGCTGTGGGAGCGGCTCCGACGAGGCTCTGCCC 1795
Qy      1741 AAGAGTGAATCTGCTCTCTTGGACCAATAATACCTACCGCTGCGCAAAATGGCTCTGTC 1800
Db      1796 CCAAGGTGAACCTGCTCTCTTGGACCAATAATACCTACCGCTGCGCAAAATGGCTCTGTC 1855
Qy      1801 TGAGCAAGGCAACCTGTAGTGTGATGAGAGAGCGAGCTGTAGCCATGGCTCCGATGAGA 1860
Db      1856 TGAGCAAGGCAACCTGTAGTGTGATGAGAGAGCGAGCTGTAGCCATGGCTCCGATGAGA 1915
Qy      1861 AAACTGTACTGTGGGCTGCAGTCTTTACCAACAGAGCTGCGGTGTGGTGGAGACA 1920
Db      1916 AAGACTGCACTGTGGGCTGCAGTCTTTACCAACAGAGCTGCGGTGTGGTGGAGACA 1975
Qy      1921 ATGGGAGGAGGCGGAGTGGCTTGGAGAGTGAAGCTTCCAGCCCTGGGCGACGAGCACT 1980
Db      1976 ATGGGAGGAGGCGGAGTGGCTTGGAGAGTGAAGCTTCCAGCCCTGGGCGACGAGCACT 2035
Qy      1981 TGTGTGGGCGCTGCTCATCTCTCTGACTGTGCTGTCTGTGACCTATTGCTTTACAG 2040
Db      2036 TGTGTGGGCGCTGCTCATCTCTCTGACTGTGCTGTCTGTGACCTATTGCTTTACAG 2095
Qy      2041 ATGACAAAATTTCAGTACTCAGACTACAGCATGTGACGCGCTTCTGGGTGCTGG 2100
Db      2096 ATGACAGAGGATTCAGTACTCAGACTACAGCATGTGACGCGCTTCTGGGTGCTGG 2155
Qy      2101 ACCAGAGAGGCGGAGTGGCTTGGGAGTGAAGCTTAAGCTTAACATCATCACCC 2160
Db      2156 ACCAGAGGCGGAGTGGCTTGGGAGTGAAGCTTAAGCTTAACATCATCACCC 2215
Qy      2161 ACCCTTCTCATATATTCAGTCTGACTAGATGATGCGCTGTGTGAGAGTGGAGAGT 2220
Db      2216 ACCCTTCTCATATATTCAGTCTGACTAGATGATGCGCTGTGTGAGAGTGGAGAGT 2275
Qy      2221 CGGTGAGTACAGACCGTGTGGCGCCCATCTGCTGCTGCTGATGCTTACCATGCTTCC 2280
Db      2276 CGGTGAGTACAGACCGTGTGGCGCCCATCTGCTGCTGCTGATGCTTACCATGCTTCC 2335
Qy      2281 CTGTGAGAGGCGATCTGGGTCAACAGCTGGGCGACACAAAAGAGGAGTACCGGAG 2340
Db      2336 CTGTGAGAGGCGATCTGGGTCAACAGCTGGGCGACACAAAAGAGGAGTACCGGAG 2395
Qy      2341 CGCTGATCTCTGACAGAGGAGTCCGTCTATCAACCAACCTGTGTAGAGACTCA 2400

```

```

Db      2396 CGCTGATCTGTCAAAAAGGATGAGATCCGCGTATCAACAGACCACTCGGAGAACCTCC 2455
Qy      2401 TGGCGAGGAGATTCACCCAGATGATGTGTGTGGTTCCTCATGTGGGAGTGTGACT 2460
Db      2456 TGGCGAGGAGATTCACCCAGATGATGTGTGTGGTTCCTCATGTGGGAGTGTGACT 2515
Qy      2461 CTGTGAGGAGTGTGATGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGT 2520
Db      2516 CTGTGAGGAGTGTGATGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGT 2575
Qy      2521 AGGCTGTGTGTGATGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGT 2580
Db      2576 AGGCTGTGTGTGATGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGT 2635
Qy      2581 CAAGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
Db      2636 CAAGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2693
Qy      2641 AGACAGCCGACCAACCAACCCGAGGATGCGCGATGCAACCTGTGATACAGAGAG 2700
Db      2694 -GCCACCAAAATGTGTACACTTGTGGGAGCGACCACTATGTCCAGAGTGTGACG- CCT 2751
Qy      2701 GAACTGACGACATTTATGCTGTGCGCTTCCCGCCCAACCAACCACTGTGACT 2760
Db      2752 GCGAGCTGAGACTGTGACCGCTGACTGTGACCAAGCGCC- CAGAAATACACTGTGACT 2810
Qy      2761 GCATCTTGTGAGCTGAGCT 2780
Db      2811 CAATCTCCAGGCTCCCAAT 2830

```

```

RESULT 7
BD274671
LOCUS   BD274671 3149 bp DNA linear PAT 17-JUL-2003
DEFINITION Matriptase, serine protease and application of the same.
ACCESSION BD274671.1 GI:33084439
VERSION JP 2002539093-A/2.
KEYWORDS Homo sapiens (human)
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

REFERENCE
1 (bases 1 to 3149)
AUTHORS Dickinson, R. B., Lin, C., Johnson, M., Wang, S., and Bnye, I.
TITLE Matriptase, serine protease and application of the same
JOURNAL Patent: JP 2002539093-A 2 19-NOV-2002;
GEORGETOWN UNIVERSITY
OS Homo sapiens (human)
PN JP 2002539093-A/2
PD 19-NOV-2002
PF 10-MAR-2000 JP 2000603721
PR 12-MAR-1999 US 60/124006
PI ROBERT B DICKSON, CHEN-YONG LIN, MICHAEL JOHNSON, SHAOWENG WANG,
PI ISTVAN ENYEDI
PC A61K45/00, A61K35/78, A61K51/00, A61P17/00, A61P35/00, PC
A61P43/00,
PC A61P43/00, C07K16/40, C12N1/15, C12N1/19, C12N1/21, C12N5/10 PC
, C12N9/64, C12N15/09,
PC C12Q1/37, G01N33/53, G01N33/577//C12P21/08, A61K49/02, C12N15/00,
PC C12N5/00
CC Matriptase, serine protease and application of the same FH
KEY FT source location/Qualifiers
FT 1.3149 /organism="Homo sapiens (human)".

```

```

FEATURES
source location/Qualifiers
1.3149 /organism="Homo sapiens (human)".

```

ORIGIN

```

Query Match 60.8%; Score 1889.8; DB 6; Length 3149;
Best local Similarity 81.4%; Pred. No. 0;

```

Matches 2222; Conservative 0; Mismatches 504; Indels 5; Gaps 3;
QY 50 GACCGCAAAACCAATGGTAGCAATGGGGCCGCAAGCCCGAGGGGGCTCTCAAGACTT 109
Db 23 GGCCTCGGGGACCAATGGGGGAGATGGGGCCGCAAGGGCGGGGGCCGAAAGACTT 82
QY 110 CGCGCGGGGACCTCAATGACAACTCCGGGCTAGAGAACTGAATGCTTTGAGGAGGGTGT 169
Db 83 CGCGCGGGGACCTCAATGACAACTCCGGGCTAGAGAACTGAATGCTTTGAGGAGGGTGT 142
QY 170 GGAGTTCCTGCTCGCAACCAATGCAAGAAAGTGAAGAACGAGCCCAAGGCGTGGT 229
Db 143 GGAGTTCCTGCTCGCAACCAATGCAAGAAAGTGAAGAACGAGCCCAAGGCGTGGT 202
QY 230 GGTGCTGTGGCAGTGTCTTCACTTCTCTTGTCTCTCCCTCATGGCTGGCTTGGTGT 289
Db 203 GGTGCTGTGGCAGTGTCTTCACTTCTCTTGTCTCTCCCTCATGGCTGGCTTGGTGT 262
QY 290 GTGGCACTTCATTATGCGAATGCGGGGTTCAAAAAGTCTCAATGGCCATCTGAGAT 349
Db 263 GTGGCACTTCATTATGCGAATGCGGGGTTCAAAAAGTCTCAATGGCCATCTGAGAT 322
QY 350 CACAATGAGATCTTTCTGATGCGTATGAGAACTCCACTCAAGATTATACGCTT 409
Db 323 CACAATGAGAAATTTTGTGATGCTTACGAACTCCACTCAAGATTATAGCTT 382
QY 410 GGCAGCAGAGTGAAGAGGCGCTGAAGCTCTGTATCAATGAAAGTCCCTGGCGGCTC 469
Db 383 GGCAGCAGAGTGAAGAGGCGCTGAAGCTCTGTATCAAGAGTCCCTGGCGGCTC 442
QY 470 CTACCAACAAAGAGCGGCTGTAACTGCTTCACTGAGAGGAGTGTCAATGCTTACTG 529
Db 443 CTACCAACAAAGAGTGTGCTGTAGAGGCTTCAAGGAGGAGGAGGCTCAATGCTTACTG 502
QY 530 GTCAAGTGTAGCATCCGCCCACTCTGCAAGAAAGAGTTGATGCGCATGCTGTGGA 589
Db 503 GTCTGAGTTCAGCATCCGCCAGCACTGTGTGAGAGAGGCGCGCTCAATGCGCAAGA 562
QY 590 GCGAGTTGTACATTTGCCACCCCGAGACCGGGCACTGAATCCTTGGTCAATCATCTGT 649
Db 563 GCGAGTTGTACATTTGCCACCCCGAGAGCGGCTGTGAGTCTTGTGTGACTTCACT 622
QY 650 GGTGGCCCTCCCAATTTGACCCCAAGATGCTCAGAGGACTCAGGACAAACAGCTCACTT 709
Db 623 GGTGGCCCTCCCAAGATTTGACCCCAAGATGCTCAGAGGACTCAGGACAAACAGCTCACTT 682
QY 710 TGCCCTGATGCCATGCTGTGAGAGAGTGAACGCTTCACTTACCTGCTTCCCAAG 769
Db 683 TGCCCTGATGCCATGCTGTGAGAGTGAACGCTTCACTTACCTGCTTCCCAAG 742
QY 770 TCCCTACCCGGGCGCATGCGCGCTGCTCAATGGGTCTGTGGGGGGAGCGCGCATCTGTGCT 829
Db 743 CCCCCTACCCCGCTCATGCGCGCTGCAATGGGCTCTGTGGGGGGAGCGCGCATCTGTGCT 802
QY 830 GAGGCTCACTCCGAAGCTTTGATGTGCTCCCTGTGATGAGCATGAGCATGCTGT 889
Db 803 GAGGCTCACTCCGAAGCTTTGATGTGCTCCCTGTGATGAGCATGAGCATGCTGT 862
QY 890 CACCGTATGATAGCTTGAAGCCCATGAGAACCCCAAGCTGTGTGGCGGTGTGGGAC 949
Db 863 CACCGTATGATAGCTTGAAGCCCATGAGAACCCCAAGCTGTGTGGCGGTGTGGGAC 922
QY 950 CTTTCACTCCCTCTCAACCTGACTTCTCTCTCCCAAGAGCTTCTCTGTCAAGGT 1009
Db 923 CTAACCTCCCTCTCAACCTGACTTCTCTCTCCCAAGAGCTTCTCTGTCAAGGT 982
QY 1010 GATAACCAATATCTGACCGGCGCAATCTGGCTTTGAGGCACTTTCTTCACTGCCCCA 1069
Db 983 GATAACCAATATCTGACCGGCGCAATCTGGCTTTGAGGCACTTTCTTCACTGCCCCA 1042
QY 1070 GATGAGCACTGTGGCGGCTTTTGTAGTGAACCCCAAGGAGCATTTAGAGCCCTCACTA 1129
Db 1043 GATGAGCACTGTGGAGGCGGCTTATGATGAAGCCCGAGGAGCATTCACAGCCCTCACTA 1102

QY 1130 TCCAGGCACTACCGGCCCAACATCACTGCACTGGAATATCAAGTGTCCCAACACCG 1189
Db 1103 CCAAGGCCACTACCCACCACTATGACTGCACTAGGAACATTTAGAGTCCCAACACCA 1162
QY 1190 GAACGTGAAGGTGCGCTTAACTCTTATCTGTGTGGAGCCCAAGTACAGAGTGGCTC 1249
Db 1163 GATGTGAAGGTGCGCTTAACTCTTATCTGTGTGGAGCCCAAGTACAGAGTGGCTC 1222
QY 1250 CTGCAACCAAGACTATGTGAGATCAACCGGGGAAAGTACTGCGGTGAGAGTGTCCAATT 1309
Db 1223 CTGCCCCAAGACTATGTGAGATCAATGGGAGAAATCTGCGGAGAGAGTGTCCAATT 1282
QY 1310 TGTGTGAGCAGACAGACAGCAAGATTAAGTCCACTTTCATTTGTATCACTGTACAC 1369
Db 1283 CGTGTCAACAGCAACAGCAACAAAGATCAAGTTCCTTCCACTCAAGATCAAGTCTACAC 1342
QY 1370 GGAACCGGGTCTCTGAGTGAAGTACCTCTCTGCACTCCCAAGACCCGCGCCAGGAT 1429
Db 1343 CGACACCGGCTCTTATGCTGAATACCTCTCTGCACTCCCAAGATCAAGTCTACAC 1402
QY 1430 GTTCATGTGCAAGACTGACCGGTGATCCGAAAGAACTGCGCTGCGACGCTGGGCGA 1489
Db 1403 GTTCATGTGCGGACCGGGGCGTGTATCCGAAAGAGCTCGCTGTGATGGTGGGCGA 1462
QY 1490 CTGCGCGGATTAATGATGAGGCTTACTGCCATGCAATGCCACCAACCAATTAAGT 1549
Db 1463 CTGCAACGACACAGCGATGAGCTCACTGCAAGTGGAGCGCGGCCCAAGTTCAGT 1522
QY 1550 CAAAAACCAATTTGCAAGCCCTCTTCTGAGCTGTGACAGTGTCAACAGACTGTGGGA 1609
Db 1523 CAAGAACAAATTTGCAAGCCCTCTTCTGAGCTGTGACAGTGTCAACAGACTGTGGGA 1582
QY 1610 CGAAGTGAACGAGAGGGCTGACAGTGTCTCTGAGAGTTTCAAGTGTTCATATGGGA 1669
Db 1583 CAACAGGAGAGAGAGGGGTGAGTGTCTCGGCCCAACCTTCAAGTGTTCATATGGGA 1642
QY 1670 GTGTCTCTTCAAGCCCAAGATGTAAATGGGAAGGAACAATGGGAGTGTGACGA 1729
Db 1643 GTGTCTCTTCAAGCCCAAGATGTAAATGGGAAGGAACAATGGGAGTGTGACGA 1702
QY 1730 GGTTCATGAGGACGCGTGAATGTCTCTTGCACCAATATATCTTACCGCTGCCAAA 1789
Db 1703 GGTCTCTCTCCCAAGGTGAACCTGTCTACCTTGTACCAAAACACTTACCGCTGCCAAA 1762
QY 1790 TGCCCTGTGTGAGCAAGGCAACCTGAGTGTGATGGGAAAGCACTGTAGCATGG 1849
Db 1763 TGCCCTGTGTGAGCAAGGCAACCTGAGTGTGATGGGAAAGCACTGTAGCATGG 1822
QY 1850 CTCCGATGAGAAAACTGTGACTGTGGGCTGTGCAATCTTTTCAAAACAGGCTCGGTGT 1909
Db 1823 CTCCGATGAGAAAACTGTGACTGTGGGCTGTGCAATCTTTTCAAAACAGGCTCGGTGT 1882
QY 1910 TGGTGGACAGAAATGCGAGCGAGGAGTGGCCCTGAGCAGGTGAGCTTCAAGCCCTGGG 1969
Db 1883 TGGTGGACAGAAATGCGAGCTGTGAGGCGAGTGGCCCTGAGCAGGTGAGCTTCAAG 1942
QY 1970 CCAGGGCCACTTGTGTGGGGCTGTGCTATCTTCTGTGACTGTGTGTCTGTGACGCTCA 2029
Db 1943 CCAGGGCCACTTGTGTGGGGCTGTGCTATCTTCTGTGACTGTGTGTCTGTGACGCTCA 2002
QY 2030 TTGCTTTAGAGATGACAAAATTTTCAAGTACTCAAGACTCAAGATGTGAGCGGCTTCT 2089
Db 2003 CTGCTTACATGATGACAAAGATTAAGTACTCAAGCCCAAGCAGTGTGAGCGGCTTCT 2062
QY 2090 GGGTGTGTGACCAAGACGAGGAGTGTCTTGGGGTGCAGAGCTGAAGCTCAACG 2149
Db 2063 GGGCTTGTGACCAAGACGAGGAGTGTCTTGGGGTGCAGAGCTGAAGCTCAACG 2122
QY 2150 TATCATCAACCAACCTTCTCTCAATGATTTCACTTGTGACTATGACTGTGCTGTGA 2209
Db 2123 CATCATCTCCACCCCTTCTTCAATGATTTCACTTGTGACTATGACTGTGCTGTGA 2182

QY	650	GGTGGCCCTTCCCATTTGACCCCGAATGCTGCAGAGACCTCAAGAACAAAGCTGCAGATT	709
Db	623	GGTGGCTTTCCCAACGGACTTCCAAAACATGACAGAGAACCCAGAACAAAGCTGCAGATT	682
QY	710	TGACCTTCAGAGCCCATGGTGCAGACAGTGAACAGCTTCACTACCTTGCTTCCCAACAG	769
QY	830	GAGCCTCAACCTTCCGAAGCTTTGATGTCGTCTCCCTGTGATGACATGACAGTGACTGTG	889
Db	803	GAGCTCAACCTTCCGAGCTTTGACCTTGGCTGTCGAGACGCGCGGAGCGAAGCTGTG	862
QY	770	TCCCTTACCGGCGCATAGCCCGCTGCAGATGGGTCCTCGCGGGGGGAGCGCGAAGCTGTG	829
Db	743	CCCTTACCCCGCTATCCCGCTGCAGATGGGCTCTCGGGGGGAGCGCGAAGCTGTG	802
QY	890	CACCGTGTGATATGCTGAGCCCTCATGAAACCCCAAGCTGTGTGTGTGTGTGTGTGTG	949
Db	863	GACGCTGACAAACCTTGAGCCCAAGAGCCCAAGCCCTGTGTGTGTGTGTGTGTGTG	922
QY	950	CTTTCACCCCTCCACAAACCTGACTTTCCTCTCTCCCAAGAGCTCTTCTTGTACAGCT	1009
Db	923	CTAACCTTCTCTTCAACCTGACCTTCACTCTCTCCCAAGAGCTCTGTATCAACACT	982
QY	1010	GATTAACCAATCTGACCGGCGCATCTGTGCTTGTAGGCCACTTCTTCTCACTGCTCCAA	1069
Db	983	GATTAACCAACTGAGCGGCGCATCTCCGCTTGTAGAGCCACTTCTTCCAGCTGCTCATG	1042
QY	1070	GATGAGCAGCTGTGGCGCTTTTGTGATGACACCCAGAGGACATTTGACAGCCCTTACTA	1129
Db	1043	GATGAGCAGCTGTGAGCGCTTACGTAAAGCCCAAGGACATTTCAACAGCCCTTACTA	1102
QY	1130	TCCAGGCCACTACCGGCCCAATCAATGACATGACATGACATGACATGACATGACATGAC	1189
Db	1103	CCAGGCCACTACCCCAATGACATGACATGACATGACATGACATGACATGACATGACATGAC	1162
QY	1190	GAACTGAAAGTGGCTTCAAACTCTTCTATCTGTGTGAACCCCAAGCTTACAGTGGGCTC	1249
Db	1163	GCATGTGAAGGTGGCTTCAAACTCTTCTATCTGTGTGAAGCCGCGCTGTGCTGGGGCAC	1222
QY	1250	CTGCACCAAGGACTTGTGTGAGATCAACGGGGGAGAAAGTACGCGTGAAGAGTCCCAATT	1309
Db	1223	CTGCCCAAGGACTTGTGTGAGATCAATGGGGGAGAAATACGCGGGAGAGAGTCCCAATT	1282
QY	1310	TGTGTGAGCAGCAACAGCAGCAAGATTTACAGTCACTTCCATTTCTGATCACTGTATCAC	1369
Db	1283	CGTCGTACACAGCAACAGATCAAGATCAAGTTCGCTTCCACTCAAGTCAAGTCTTACAC	1342
QY	1370	GGACACCGGGTTCATAGTGTGATACCTCTCTAGACTCCAAAGAACCCGATGCCAGGAGAT	1429
Db	1343	CGACACCGGGTTCATAGTGTGATACCTCTCTAGACTCCAAAGAACCCGATGCCAGGAGAT	1402
QY	1430	GTTTCAATGTGCAAGACTGTGACCGGTGCATCCGAAGAACTGTGCTGTGCAAGCTGTGGGACA	1489
Db	1403	GTTTCAATGTGCGCAAGGGCGGTGTATCCGAAGAGACTGTGCTGTGCTGTGGGCGCA	1462
QY	1490	CTGCCCGATTATGATGATGACCTTACTCTGCTATGCAATCCACCCACAGTTTCAAGTG	1549
Db	1463	CTGCACCGCACACAGAGATGAGCTCAACTGTGCAAGCTGTGCAAGCTGTGCAAGCTGTG	1522
QY	1550	CAAAAAACAGTTCTGTGAACCCCTCTTCTGGGCTGTGTACAGTGTCAACGACTGTGGAGA	1609
Db	1523	CAGGAACAGTTCTGTGAACCCCTCTTCTGGGCTGTGTGAAGTGTGAACGACTGTGGAGA	1582
QY	1610	CGGAAGTGAACAGAGAGGCTGTGAGCTGTCTGTGGAGTTTCAAGTGTTCATATGGGAA	1669
Db	1583	CAACAGCAGCAGAGAGAGGCTGTGAGTGTCTGTGGAGCTTCAAGTGTTCATATGGGAA	1642
QY	1670	GTGTCTCTCTAGAGCAGAAAGTGAATGGGAAGCAACTGTGAGATGGAGTCTGACGA	1729
Db	1643	GTGTCTCTCTAGAGCAGAAAGTGAATGGGAAGCAACTGTGTGGAGCAGGCTCTGACGA	1702

QY	1730	GGCCTTCAATGACAGCGGTGAATGTCGCTCTTGTGACCAAAATATCTACCGCTGGCAAAA	1789
Db	1703	GGCCTCTCGCCCCAAGGTGAACGTGTGCATCTTGTAACAAACACACTTACCGCTGCTTCAA	1762
QY	1790	TGGCCTCTGTCTGAGCAAGGCAACCCCTGAATGTGTATGGGAAGACGACTGTAGCGATGG	1849
Db	1763	TGGGCTCTGCTTGAAGCAAAAGGCAACCCCTGAATGTGTACCGGAAGGAGCACTGTAGCGAGG	1822
QY	1850	CTCCGATGAGAAAACTGTGACTGTGGGCTTCGCAATCCCTTACCAACAGGCTCGGTGTGGT	1909
Db	1823	CTCAGATAGAGAAGCACTCGCACTGTGGGCTCGGTCATTCACGAGACAGGCTCGTGTGT	1882
QY	1910	TGATGGCAAGATGGGAGCGAGAGGGGCAATGGCCCTGGAGAGGTAGGACTTCAAGCCCTGGG	1965
Db	1883	TGGGGGCAACGATTCGGATGTAGGGGGAATGGGCCCTGGAGAGGTAGGACTTCTGGG	1944
QY	1970	CCAGGGCCCATTTGTGTGGGGCCCTGGCTCATCTTCTCTGAATGTGGCTGTCTGTGCAAGCTCA	2022
Db	1943	CCAGGGCCCATCTGGGGGTGCTTCCCTCATCTCTCCAACTGGCGTGTCTGTGCGCA	2007
QY	2030	TTGCTTTAGAGATGACAAAATTTCAAGTACCTCAGACTACACGATGTGACAGGCTTCTCT	2088
Db	2003	CTGTCAATCGATGTACAGAGATTCAGGTACTCAGACCCCAAGCAAGTGAAGGCTTCTCT	2066
QY	2090	GGGTCTGTGACACAGACCAAGCCGAGTGGCTCTGGGGGTGCAAGACTGAAGCTCAACG	2144
Db	2063	GGGCTTGACAGACACAGACCCAGCCGAGCGCCCTGTGGGTGCAAGGAGCCGAGGCTCAAGG	2122
QY	2150	TATCATCAACCACTTCTCTTCAATGATTTCACTTGCATATGACATGACATGACCTTCTGGA	2200
Db	2123	CATCATCTTCCACCCCTTCTTCAATGACTTCACTTTCAGCTTATGACATGACATGACCTTCTGGA	2188
QY	2210	GCTGAGAGAATGCGGTGAGATACAGACCGTCTGGCGCCCATCTGTCTGCTGATGTACTAC	2266
Db	2183	GCTGAGAGAAACCGGAGAGTACAGCTCCATGTGTGGGCCCATCTGCTGCGCGAGCCCTC	2244
QY	2270	CCATGCTTTTCCCTGCTGTGGCAAGGCCATCTGGGTTCACAGGCTGGGGGACACAAAAGAGGG	2332
Db	2243	CCATGCTTTTCCCTGCTGTGGCAAGGCCATCTGGGTTCACAGGCTGGGGGACACACCAGATGG	2300
QY	2330	AGGTACCCGAGCGCTGATCCTGCAAGAGGGGTAGATCCCTGTATCAACCAAGACCACTTG	2388
Db	2303	AGGCACTGGCGGCTGTATCTTCGAAAAGGGTATGAGATCCGGTCAATCAACCAAGACCACTTG	2366
QY	2390	TGAGGACCTCATGTCGCGAGACGATCAACCCACGATGATGTGTGTGGGTTTCTCTCAGTGG	2444
Db	2363	CGAGAACCTCTGTGCGGCAACATCAACGCGGCAAGATGTGTGTGGGTTTCTCTCAGCGG	2422
QY	2450	GGGTGTGACCTCTGTCAAGGTGACCTGTGTGGCCCTTGTCAAGGCGGAGAAAATGG	2500
Db	2423	CGGCGTGACTCTCTGTCAAGGTGATTCGAGGGGAGCCCTGTCAAGGCTGGAGGCGGATGG	2488
QY	2510	GCGAATGTTCCAAGCTGTGTGTGTGAGCTGGGGGTGAAGGCTGCTCAAGGAAACAAGCC	2566
Db	2483	GCGAATCTTCAAGGCGGATGTGTGTGTGAGCTGGGGGAGACGCTGTGCTCAGAGGAAACAAGCC	2544
QY	2570	AGGCTGTGTACAAAGGCTCCCTGTAACTTGTGGGACTGGATTCAAAGAGACACTGGGGTATA	2622
Db	2543	AGGCTGTGTACAAAGGCTCCCTGTGTGTGGGACTGGATTCAAAGAGACACTGGGGTATA	2600
QY	2630	GCAGCATGACAGACAGCGACCAACAACCCACAGGATGGCTGACATGCAACACTGG	2688
Db	2603	GGGGCCGGG---GCCAACCAATGTGTACCTGCGGGGCAACCCATCGTCCACCCAGT	2656
QY	2690	ATACAGAGAGAGAACTGACGACATTTATGTGTGTGGCTTCCCCCCCCCAACAACCCA	2744
Db	2660	GTCACAG-CCTGACAGCTGTGAGACTGGAACGCTGATCTGCAACAGCGCC-CAGAACATA	2711
QY	2750	GACTGTGAATGCTATCCTTGAAGCTAGAGT	2780
Db	2718	CACGTGAATCAATCTCCAGGGCTCCAAT	2748

```

RESULT 9
AF133086 3120 bp mRNA linear PRI 01-OCT-1999
LOCUS Homo sapiens membrane-type serine protease 1 mRNA, complete cds.
DEFINITION AF133086
ACCESSION AF133086.1 GI:6002713
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Takeuchi,T., Shuman,M.A. and Craik,C.S.
  Reverse biochemistry: use of macromolecular protease inhibitors to
  dissect complex biological processes and identify a membrane-type
  serine protease in epithelial cancer and normal tissue
  Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11054-11061 (1999)
JOURNAL MEDLINE 99432178
PUBMED 10500122
REFERENCE
2 (bases 1 to 3120)
  Takeuchi,T., Shuman,M.A. and Craik,C.S.
  Direct Submission
  Submitted (04-MAR-1999) Dept. Pharm. Chem., University of
  California, San Francisco, 513 Parnassus Ave., Box 0446, San
  Francisco, CA 94143-0446, USA
JOURNAL Location/Qualifiers
FEATURES
source
1..3120
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /cell_line="PC-3"
  /cell_type="prostate carcinoma"
  /note="WT-Sp1; chymotrypsin clan serine protease"
  /codon_start=1
  /product="membrane-type serine protease 1"
  /protein_id="AAF00109.1"
  /db_xref="GI:6002714"
  /translation="MGSDPRKGGGPGKFGALKYNSHEKNGLEGEVPLPVNV
  KVEKHGPERWVLAVALIGLLVLLGIFLWMLQYRVRVQKFNMGRLTNFNF
  DAVENSHGPERWVLAVALIGLLVLLGIFLWMLQYRVRVQKFNMGRLTNFNF
  IPOHVEAEERVAERVYMLPPRARSLSKPVTSVSVSTKTVORPQSDSCPL
  HARVELEMRFTGPPDSDPYPAHARCOMALRGDADSVLSITRSDLSADERSDVL
  TVYNTLSMPEHALYVOLCTGYTPSYNLTTHSQNVLLTLTNTERRHGFATFOL
  PRMSCGSRKRAQGTENSPLYGHYPNITDCTMNI EYNNQHVAFRFEFLYEPV
  PAGCPKDYVEINTEGRKYGERSQFVVTISNKTIVRPHSDQYDTGFAEYLSYDS
  DPCPQFCTCRCTRKEIRKCDGMADCTDSDHSCNCDAGQPTCKXKFCPLFWC
  DSVNDCGNSDEQSCPAQCFRCSNGKLSKQOCCNGKDDCGSDSDASCPKXVVT
  CTKHTRCLNGCLSKNGPECDGKEDCSGDEKDCDGLRFTFOARVGGTADBG
  EMPQVYSALHAGGHIICGASLISPMVLVSAHCYIDDGFRSDDTPTWATFGLHDQS
  ORSAGVDERLRKRLIISHFEFNDFTFDYDALLELEKPAEYSSMVRPCLPDASHVFP
  AGKAIWVGWHTYGGTGALILQGEIRVINVOTTCENLIPQITPRMWCFTSLGSG
  DSCGDSGGPLSSVEADRIFGAGVAVSWGDGCAQRNKPQVYTRPLFRDWMKENTGV"
ORIGIN
Query Match 60.7%; Score 1884.8; DB 9; Length 3120;
Best Local Similarity 82.8%; Pred. No. 0;
Matches 2164; Conservative 0; Mismatches 447; Indels 1; Gaps 1;

```

```

Db 188 GCCCGGGGCGCTGGGTGTCGAGCCGTCGATCGGCTCTTGCTTGTCTGG 247
Qy 274 TGCTGGCTTGCTGGTGTGGACCTTCATATGGMAATGGCGGTTCAAAAAGTCTTCA 333
Db 248 GGATCGGCTTCTGGTGTGGCATTTGTGATACCGGAGCGTGGCTTCAAGGTCCTTCA 307
Qy 334 ATGGCCATCTGAGGANTACAAATAGATCTTTCTGATGGTATGAGACTCCACTCCA 393
Db 308 ATGGCTACATGAGGATACAAATGAGAAATTTTGTGATGCTTACAGAACTCCACTCCA 367
Qy 394 CAGAGTTATACCTGCGCCAGCCAGGTGAGAGAGCGGTGAAGCTGCTGTACANTAG 453
Db 368 CTGAGTTGTAACTGGCCAGAGAGGTGAGAGCGGCTGAAGCTGCTGTACAGCGGAG 427
Qy 454 TCCCTGCTGGGTGCTCCCTACCAAGAGTGGCTGTAACTGCTTGTAGAGGCGAGT 513
Db 428 TCCCATTTCTGGGCCCCCTTACCAAGAGAGTGGCTGTAGAGGCGCTTCAAGAGGCGAG 487
Qy 514 TCATGCTACTACTGTCAGTGAAGTTCAGATCCGCCACACTGGCAGAAAGGTTGATC 573
Db 488 TCATGCTACTACTGTCAGTGAAGTTCAGATCCGCCACACTGGTGAAGAGCGGAGC 547
Qy 574 GCGCCATGCTGTGAGAGCGAGTTGTAACATTTGCCACCCGAGCAGCGCACTGAATCT 633
Db 548 GCGTCATGGCCAGAGCGCGGTATGATGCTGCCCCGCGGCGGCTCCCTGAAGTCT 607
Qy 634 TCGTGTAACTACTGTGTGGCTTCCCATTTGACCCCGAAGATGCTGCAGAGATCTAG 633
Db 608 TTGTGTACTCTGATGTGGCTTTTCCACGAGATCTCAAAACAGTACAGAGACCCAG 667
Qy 694 ACAACAGCTGCAAGTTTGGCTGCAATGCCATGTGTCAGACAGTACAGCTTCACTTACC 753
Db 668 ACAACAGCTGCAAGCTTGGCTGCAATGCCGAGGTGTGAGAGCTGATGCTTCACTTACC 727
Qy 754 CTGGTCTCCCAACAGTCTTACCCGAGCATATGCCGCTGTCAGATGGTCTTGGCGGAG 813
Db 728 CGGGCTTCCCTACAGCCCTTACCCGCTCATGCGCGCTGACAGTGGCGCTTGGCGGAG 787
Qy 814 AGCGGACCTGTGTGAGGCTTCACTTCCGAGCTTGTATGCTGCTCCCTGTGATGAGC 873
Db 788 AGCGGACCTGTGTGAGGCTTCACTTCCGAGCTTGTGATGCTGCTCCCTGTGATGAGC 847
Qy 874 ATGGCAGTGAACCTGTGACCTGTGTATGATAGCTGAGCCCATGGAACCCACGCTGTGG 933
Db 848 GGGGAGGAGCACTGTGTAGAGGTGTAACAACCTGAGCCCATGAGCCCAAGCCCTTGG 907
Qy 934 TGGCGCTGTGTGACCTTCTTCACTTCTTCAACCTGAATTTCTTCTTCCAGAACG 993
Db 908 TGCAGTTGTGTGACCTTCACTTCTTCTTCAACCTGAATTTCTTCTTCCAGAACG 967
Qy 994 TCTTCTTGTACAGCTGTATACCTATCTGACCCGCGCATCTCTGCTTTGAGGCACTT 1053
Db 968 TCTTCTTGTACAGCTGTATACCTATCTGACCCGCGCATCTCTGCTTTGAGGCACTT 1027
Qy 1054 TCTTCAAGCTGCGCAAGATGAGCACTGTGCGGCTTTTGTAGTACACCCAGGAGCAT 1113
Db 1028 TCTTCAAGCTGCGTATGATGAGCACTGTGAGCGGCTTGTAGTACACCCAGGAGCAT 1087
Qy 1114 TTAGAGCCCTTATCTATCCAGGCACTTACCGGCCCAATCACTCACTGACATGAAATATCA 1173
Db 1088 TCAAGAGCCCTTATCTATCCAGGCACTTACCGGCCCAATCACTGACATGAAATATCA 1147
Qy 1174 AGTGGCCCAACACCGAGATGAGAGGCGCTTCAACTTCTATCTGTGTGAGCCCA 1233
Db 1148 AGTGGCCCAACACCGATGTGAGAGGCGCTTCAACTTCTATCTGTGTGAGCCCG 1207
Qy 1234 ACGTACAGTGGGCTCTGTGACCAAGGCTATGTGTGATCAACGCGGAGAGATCTGCG 1293
Db 1208 GCGTGGCTGCGGAGCACTGCGCCCAAGGACTACGTGTGAGATCAATGGGAGAAATCTGCG 1267
Qy 1294 GTGAGAGGTCCTCAATTTGTGTGAGCAACACAGCAGCAATTAAGTCCATTTCCATT 1353
Db 1268 GAGAGAGGTCCTCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1327

```

QY 1354 CTGATCACTGTAACAGGACACCGGCTTCTAGCTAGTAAGTACTCTCTTACGACTCCAGG 1413
 DB 1328 CAGATCAGTCTTACACCGACACCGGCTTCTTACTGATAATACCTCTCTGACTCCAGTG 1387
 QY 1414 ACCCGTCCAGGAGATGTTCAATGACGAAAGCTGACGAGTGCATCCGAAAGAACTGGCT 1473
 DB 1388 ACCCATGCCCCGGGACGTTTACGTGCGGACGGGGCGGTGTATCCGGAAAGACTGGCT 1447
 QY 1474 GCGACGCGTGGGGAGACTGCCCCGGAATTATAGTATGAGCGTTTACTGCGATGCAATGCCA 1533
 DB 1448 GTGATGGCTGGGCGACTGACCGACCGACAGCATGAGTCAACTGCAATGGGACGCGC 1507
 QY 1534 CCCACCACTTCAAGTGGAAAAACGATTTCTGCAAGCCCCCTCTCTGGGTCTGAGCAAGTG 1593
 DB 1508 GCCACCAAGTTCAGGTGCAAAACAAAGTTTGCAGAGCCCTCTCTGGGCTTGGGACAGTG 1567
 QY 1594 TCAACGACTGTGGGGAGCGAAAGTGAACGAGGAGGCTGCACCTGTCTGTGGAGTTTCA 1653
 DB 1568 TGAACGACTGGGAGACACAGGACGACGAGGGGTGCATGTGTCGGCCCAAGCCTTCA 1627
 QY 1654 AGTGTTCATGGGAAAGTGTCTCTCTGACGACCGAAGTGTATGGGAAGGACACTGTG 1713
 DB 1628 GGTGTTCATGGGAAAGTGTCTCTGAAAGCCAGAGTGCATGGGAGGACGACTGTG 1687
 QY 1714 GAGATGGGTCTGACGAGAGCTTCAATGTGACAGCGGTGAATGTCTGTGACCAAAATATA 1773
 DB 1688 GGGACGGGCTCGAGAGGCTCTCTGCCCCCAAGGTGAAGTGTGACTTGTACCAACACA 1747
 QY 1774 CCTACCGCTGCGAATAATGCTCTGTCTGACGACGAGGCAACCTGAGTGTATGGGAGA 1833
 DB 1748 CCTACCGCTGCTCATATGGGCTCTGTGTGACGAAAGGCAACCTGAGTGTGACGGAGG 1807
 QY 1834 CGGACTGTAGGATGCTCTCGATGAGAAAACTGTGACTGTGGCTGATCCTTTACCA 1893
 DB 1808 AGGACTGTAGGAGCGGCTCAGATGAGAAAGACTGCTGAGCTGTGGCTCATTTACGA 1867
 QY 1894 AACAGGCTCGGCTGTGTGTGTGACGAAATGCGGACGAGGAGTGGCCCTGGGAGGTGA 1953
 DB 1868 GACAGGCTCGGTGTGTGTGTGGGACGAGATGCGGATGAGGGGAGTGGCCCTGGGAGGTGA 1927
 QY 1954 GCCCTCAAGCCCTGAGGACGAGGACCACTTGTGTGGGAGCTGTGCTCATCTCTGACTGGC 2013
 DB 1928 GCTTGCATGCTCTGTGGCCAGGGCCACATGTGGGTGCTTCCCTCATCTCTCCCAACTGGC 1987
 QY 2014 TGGTCTCTGAGCTCATTTGCTTTTCAAGATGACAAAAATTTCAAGTACTCAAGTACAGA 2073
 DB 1988 TGGTCTCTGCGGACACTGTCTACATCATGATGACAGAGATTCAAGTACTCAAGCCCCAGC 2047
 QY 2074 TGTGAGCGGCTTCTGTGGGTCTGTCTGACGACGAGAGCAAGCGAGTGCCTGTGGGGTGCAGG 2133
 DB 2048 AGTGAACGCGCTTCTGTGGCTTGTGACGACGAGGACGAGCGCCCTGGGAGTGCAGG 2107
 QY 2134 AGCTGAAGCTCAAAAGTATCATCAACCCACCTTCTCTTCAATGATTTTCACTTGCATATG 2193
 DB 2108 AGCGGAGGCTCAAGGACTCATCTCTCCACCTTCTTCAATGATTTTCACTTGCATATG 2167
 QY 2194 ACATGCTTGTGTGAGCTGAGAAAGTGTGTGAGTACGACCGGTCTGTGGCCCTCATCT 2253
 DB 2168 ACATGCTTGTGTGAGCTGAGAAAGTGTGTGAGTACGATGATGATGATGATGATGATGAT 2227
 QY 2254 GCTTCCCTGATGCTTCCATGATGCTTCCCTGCTGGGACGAGGCTGTGGGTCAAGAGCTGGG 2313
 DB 2228 GCTTCCCTGAGCTCTTCCATGATGCTTCCCTGCTGGGACGAGGCTGTGGGTCAAGAGCTGGG 2287
 QY 2314 GGCACACAAAAGAGGAGGTACCGGAGCGCTGATCTTGCAGAAAGGTGATCCGTGTCA 2373
 DB 2288 GACACACCAAGATGAGGAGCACTGGGCGCTGATCTTGCAGAAAGGTGATCCGAGTCA 2347
 QY 2374 TCACACGACGACCTGTGTGAGACCTTCAATCCGAGAGCATCAACCCACGAAATGATGTGTG 2433
 DB 2348 TCACACGACGACCTGTGTGAGACCTTCAATCCGAGAGCATCAACCCACGAGATGATGTGTG 2407

QY 2434 TGGGTTCTCTAGTGGGAGTGTGAGACTCTCCGACGGGTGACTGTGAGCCCTTGTCAA 2493
 DB 2408 TGGGCTTCTCTAGCGGCGGCGGTGAGCTCTGCGCAGGATGATTCGGGGGACCCCTGTCCA 2467
 QY 2494 GCGCGGAGAAAGATGGGCGCAATTTCCAGGCTGTGTGTGACTGGGAGAGGCTGGC 2553
 DB 2468 GCGTGAAGCGGAGTGGCGGATCTTCCAGGCCGTGTGTGAGCTGGGAGAGCGCTGGC 2527
 QY 2554 CTCGAGAGAACCAAGCCAGGCGGTGACACAAAGCTCTGTGATTCGGGACTGTGATCAAG 2613
 DB 2528 CTCGAGAGAACCAAGCCAGGCGGTGACACAAAGCTCTGTGATTCGGGACTGTGATCAAG 2587
 QY 2614 AGCACACTGGGATATAGCAGCATGACAGACA 2645
 DB 2588 AGAACACTGGGATATAGGGGCGCGGACACCA 2619

RESULT 10
 AB030036
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens
 Homo sapiens (human)
 Bujakytora; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 (sites)
 Yamaguchi, N. and Mitsui, S.
 Molecular cloning of a novel transmembrane serine protease
 expressed in human prostate
 Unpublished
 2 (bases 1 to 3128)
 Yamaguchi, N. and Mitsui, S.
 Direct Submission
 Submitted (14-JUL-1999) Nozomi Yamaguchi, Kyoto Prefectural
 University of Medicine, Res. Ins. Geriatrics; Kawarachi Hirokoji,
 Kyoto, Kyoto 602-8566, Japan (E-mail: nozomi@koto.kpu-m.ac.jp,
 Tel:81-75-5848, Fax:81-75-251-5848)

FEATURES
 source
 CDS

1..3128
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /feature_type="prostate"
 16..2583
 /codon_start=1
 /product="prolactamin"
 /protein_id="BAB20376.1"
 /db_xref="gi:12249015"
 /translation="MGSDRARKGGGPKDPGAGLKYNRHEKNGLEGEVFLPVNV
 KKEKHGGRVVAALVILGLVLGIFLVMHLQYDVAOVQFNGYMBITNENFV
 DAYNSNTSEFVSLASVKYDKLKLISVPEPLGPKHESAVTAFESESVIAYVSEFS
 DVAENSTGRLKAGGTENSPPYPHYPPNIDCTNVIENVNQHVAERFRFYLLBPGV
 IPQHLVERAEVVAEBRYVMTPPARSLCQALGADSDVSLIFRSDDLSCBERGSDIV
 TVYNLISPMEBHALVOLCGTYPSPNLTFFHSSQNVLLITLTNERRRPGEBATFOL
 PRMSSCGRLKRAQGTENSPPYPHYPPNIDCTNVIENVNQHVAERFRFYLLBPGV
 PAGCPQDQYVINGEKYCGERSQPVYNSNKKITVRFSQDSYDTGFLAELSYDS
 DPCHGPQPCRRGRICRKEIRCDGMADCTDHSDELNCSGDAHQFTCKNPKFLFWC
 DSVNDGDNSEBOGSCPAQTPRCNSNKKLSKQCCNCKDDCGSDSASCPKVVVT
 CTNHYRCLNLCISKNPEDCGKEDCSDEKDDCDGLASFTRQAVVGGTDDG
 EMPQVSLHALGQGHICGASLISPMVLVAACHYDDGFRSDTQMTVFLGLHDS
 QRSAPGVQERLKRILISHPFNDPFPFDIDALIELEKPAEYSWVRPCLPDAHVFP
 AGKAIWVGWHTQYGGTALILQKGIIRVINTTCENMLILQQLTPRMKCVGLSGV
 DSCGDSGSLSSVADGRIRFQAGVSVSGDCAQGNKRGVYTRLEPLFDWIKENGV"

ORIGIN
 Query Match 60.7%; Score 1884.6; DB 9; Length 3128;
 Best Local Similarity 81.4%; Pred. No. 0;
 Matches 2222; Conservative 0; Mismatches 504; Indels 5; Gaps 3;

50 GACCGCAAAACCATGGGTAGCAATCGGGCCGCAAGCGCGGAGGGGCTCTCAGGACTT 109

```

Db      3  GGCCTCGGGGACATGGGGAGCGATCGGGCCCGAAGGCGAGGGGCGCCGAAAGACTT 62
Qy      110 CGGCGCGGACTCAAGTACAACTCCCGCTAGAGAACATGAATGGCTTTGAGAGGGTGT 169
Db      63  CGGCGCGGACTCAAGTACAACTCCCGCAAGAGAGTGAATGGCTTTGAGAGAGCGGT 122
Qy      170  GGAATTCCTGCTCCGCAACAATGCGCAGAGAAAGTGAAGAGCGAGCGCCCAAGCGCTGGT 229
Db      123  GGAATTCCTGCTCCGCAACAAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 182
Qy      230  GGTGCTGAGTGGAGTGTCTGTTCACTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 289
Db      183  GGTGCTGAGTGGAGTGTCTGTTCACTTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
Qy      290  GTGGCACTTCATTTATGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCGAGAT 349
Db      243  GTGGCACTTCATTTATGGAATGTGCGGGTTCAAAAAGTCTTCAATGGCCATCGAGAT 302
Qy      350  CAGAAATGAGATCTTTTGTGATGCTATGAGAACTTCACTTCACTTCACTTCACTTCACT 409
Db      303  CAGAAATGAGATCTTTTGTGATGCTATGAGAACTTCACTTCACTTCACTTCACTTCACT 362
Qy      410  GGCACCCAGTGAAGAGGCGCTGAAGCTGTGTAACAATGAAGTCCCTGCTCTGAGTCC 469
Db      363  GGCACCCAGTGAAGAGGCGCTGAAGCTGTGTAACAAGTCCCTGCTCTGAGTCC 422
Qy      470  CTACCAAGAAAGTGGCTGTGTAAGTCTGCTTCACTGAGGAGTGTATCGCTACTACTG 529
Db      423  CTACCAAGAAAGTGGCTGTGTAAGTCTGCTTCACTGAGGAGTGTATCGCTACTACTG 482
Qy      530  GTGAGGTTGAGATCTTCCCGCAACTGCGCAGAGAGGTTGATGCGGCAATGGCTGTGA 589
Db      483  GTGAGGTTGAGATCTTCCCGCAACTGCGCAGAGAGGTTGATGCGGCAATGGCTGTGA 542
Qy      590  GCGAGTTGTAACATTTSCACCCCGAGCAGCGGCACTGAATCCCTGCTGTAACATCTGT 649
Db      543  GCGAGTTGTAACATTTSCACCCCGAGCAGCGGCACTGAATCCCTGCTGTAACATCTGT 602
Qy      650  GGTGCTTCCCGCAACTGCGCAGAGGTTGATGCGGCAATGGCTGTGA 709
Db      603  GGTGCTTCCCGCAACTGCGCAGAGGTTGATGCGGCAATGGCTGTGA 662
Qy      710  TGGCCCTGCAATGCGCAGAGGTTGATGCGGCAATGGCTGTGA 769
Db      663  TGGCCCTGCAATGCGCAGAGGTTGATGCGGCAATGGCTGTGA 722
Qy      770  TCCCTACCCGCGCATGCGCAGAGGTTGATGCGGCAATGGCTGTGA 829
Db      723  TCCCTACCCGCGCATGCGCAGAGGTTGATGCGGCAATGGCTGTGA 782
Qy      830  GAGCTTCACTTCCGCAAGTGTGCTCTCTGTAAGCAATGGCAGTGAAGTGT 889
Db      783  GAGCTTCACTTCCGCAAGTGTGCTCTCTGTAAGCAATGGCAGTGAAGTGT 842
Qy      890  CACCGTGAATGATTAACCTGAGAGCCCATGAGAACCCCACTGTGTGTGTGTGTGTGTGT 949
Db      843  CACCGTGAATGATTAACCTGAGAGCCCATGAGAACCCCACTGTGTGTGTGTGTGTGT 902
Qy      950  CTTCTCACTTCCGCAAGTGTGCTCTCTGTAAGCAATGGCAGTGAAGTGT 1009
Db      903  CTTCTCACTTCCGCAAGTGTGCTCTCTGTAAGCAATGGCAGTGAAGTGT 962
Qy      1010  GATTAACCAATGATGAGCGGCAATCTGCTTTGAGGCACTTTCTTCAAGTGTCCCA 1069
Db      963  GATTAACCAATGATGAGCGGCAATCTGCTTTGAGGCACTTTCTTCAAGTGTCCCA 1022
Qy      1070  GATGAGCACTGTGCGGCTTTTGTGAGTGAACCCCAAGGCAATTTAGCAGCCCTACTA 1129
Db      1023  GATGAGCACTGTGCGGCTTTTGTGAGTGAACCCCAAGGCAATTTAGCAGCCCTACTA 1082
Qy      1130  TCAGGCACTACCGGCAATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1189

```

```

Db      1083  CCGAGGCACTACCCACCAACATTTGATGATGAGAACATGAGGTGCCCAACAACA 1142
Qy      1190  GAAGTGAAGTGTGCTTCAAACTTCTATCTGTGTGAGACCCCAAGTATGAGGCTC 1249
Db      1143  GAATGTGAAGTGTGCTTCAAACTTCTATCTGTGTGAGACCCCAAGTATGAGGCTC 1202
Qy      1250  CTGACCAAGCACTATGTGAGATCAACGGGGAGAAATGATCGGTGTGAGTGTCCAGT 1309
Db      1203  CTGACCAAGCACTATGTGAGATCAACGGGGAGAAATGATCGGTGTGAGTGTCCAGT 1262
Qy      1310  TGTGTGAGAGCAACAGCAGCAAGATTACAGTCACTTCCATTTCTGATCACTGTAC 1369
Db      1263  TGTGTGAGAGCAACAGCAGCAAGATTACAGTCACTTCCATTTCTGATCACTGTAC 1322
Qy      1370  GAGACCGGCTTCTGAGTGTGATCTTCTTACAGTCTTCAAGCAAGCCGTTGCCAGGAT 1429
Db      1323  GAGACCGGCTTCTGAGTGTGATCTTCTTACAGTCTTCAAGCAAGCCGTTGCCAGGAT 1382
Qy      1430  GTTCATGTGCAAGCTGGAAGTGTGATCGGAAAGGAACTGCGCTGCAAGCGCTGAGCA 1489
Db      1383  GTTCATGTGCAAGCTGGAAGTGTGATCGGAAAGGAACTGCGCTGCAAGCGCTGAGCA 1442
Qy      1490  CTGCGGCAATTAATGATGAGGTTACTGCGATGCAATGCAACCAAGTTCAGTGT 1549
Db      1443  CTGCGGCAATTAATGATGAGGTTACTGCGATGCAATGCAACCAAGTTCAGTGT 1502
Qy      1550  CAAAAACAGTTCGTGAAGCCCTCTTGTGGTCTGTGACAGTGTCAACGATGTGGGA 1609
Db      1503  CAAAAACAGTTCGTGAAGCCCTCTTGTGGTCTGTGACAGTGTCAACGATGTGGGA 1562
Qy      1610  CCGAATGACAGAGAGGCTGACAGTGTCTGTGGAAGTTTCAAGTTCGAATGGGA 1669
Db      1563  CCGAATGACAGAGAGGCTGACAGTGTCTGTGGAAGTTTCAAGTTCGAATGGGA 1622
Qy      1670  GTGTCTCTCCAGAGCCAGAGTGTATGGAAGGCAACTGTGAGATGGTCTGACGA 1729
Db      1623  GTGTCTCTCCAGAGCCAGAGTGTATGGAAGGCAACTGTGAGATGGTCTGACGA 1682
Qy      1730  GGTTCATGTGACAGTGTGATGTGTCTCTTGCACCAATATTACTTACCGCTGCCAAA 1789
Db      1683  GGTTCATGTGACAGTGTGATGTGTCTCTTGCACCAATATTACTTACCGCTGCCAAA 1742
Qy      1790  TGGCTCTGTGACAGAGGCAACCTGAGTGTATGGAAGGCAAGTGTGAGTGT 1849
Db      1743  TGGCTCTGTGACAGAGGCAACCTGAGTGTATGGAAGGCAAGTGTGAGTGT 1802
Qy      1850  CTCGATGAGAAAACTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1909
Db      1803  CTCGATGAGAAAACTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1862
Qy      1910  TGGTGGCAAGAAATGGCGACAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1969
Db      1863  TGGTGGCAAGAAATGGCGACAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1922
Qy      1970  CAGGCGCACTTGTGTGAGGCTCGCTCATCTCTCTGACGAGGCTGTGAGTGT 2029
Db      1923  CAGGCGCACTTGTGTGAGGCTCGCTCATCTCTCTGACGAGGCTGTGAGTGT 1982
Qy      2030  TTGCTTCAAGTGAAGAAAAATTTCAAGTACTCAGACTTCAAGATGTGAGCGGCTTCT 2089
Db      1983  TTGCTTCAAGTGAAGAAAAATTTCAAGTACTCAGACTTCAAGATGTGAGCGGCTTCT 2042
Qy      2090  GGGTGTGTGACAGAGCAAGAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2149
Db      2043  GGGTGTGTGACAGAGCAAGAGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2102
Qy      2150  TATCATCAACCACTTCTCTCAATGATTTCACTTCACTTCACTTCACTTCACTTCACT 2209
Db      2103  TATCATCAACCACTTCTCTCAATGATTTCACTTCACTTCACTTCACTTCACTTCACT 2162
Qy      2210  GCTGAGAAAGTGTGTGAGTACAGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2269
Db      2163  GCTGAGAAAGTGTGTGAGTACAGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2222

```

QY 2270 CCATGTCCTCCCTGCGGAGGCCATCTGGGTCAAGGCTGGGGGACACAAAGAGG 2329
DB 2223 CCATGTCCTCCCTGCGGAGGCCATCTGGGTCAAGGCTGGGGGACACAAAGAGG 2282
QY 2330 AGGTACCGGAGGCGGTATCTGTGAGAGGGTGTAGATCCGTGATCAACAGACCACTG 2389
DB 2283 AGGCACTGGGCGGCTATCTGTGAGAGGGTGTAGATCCGTGATCAACAGACCACTG 2342
QY 2390 TGAGACCTCATATCCGACAGAGATCAACCCAGATGATGTGTGGGTTTCTCATGTG 2449
DB 2343 CGAGAACTCTCTCCGAGAGATCAACGCGGCAATGATGTGTGGGTTTCTCATGTG 2402
QY 2450 GGGTGTGAGACTCTGCGGAGGATGATCTGTGAGGCTCCCTGTTCAGAGCGGAGAAATG 2509
DB 2403 CGGCGTGAATCTCTGCGGAGGATGATCTGTGAGGCTCCCTGTTCAGAGCGGAGAAATG 2462
QY 2510 GCGAATGTTCCAGGCTGTGTGTGAGGCTGGGGTGAAGGCTGCGCTCAGAGAAACAAGCC 2569
DB 2463 GCGGATCTTCCAGGCGGCTGTGTGTGAGGCTGGGGTGAAGGCTGCGCTCAGAGAAACAAGCC 2522
QY 2570 AGGCGTGTACACAAGGCTCCCTGTGATGTGTGGGACTGGATCAAGAGCACTGGGGTATA 2629
DB 2523 AGGCGTGTACACAAGGCTCCCTGTGATGTGTGGGACTGGATCAAGAGCACTGGGGTATA 2582
QY 2630 GCGACATGACACAGACGCCACCAACAACCCAGAGGATGCCGACATGACACCTGG 2689
DB 2583 GGGGCGCGGG---GCCACCCAAATGTGTACACTGCGGGGCCCAACCATGTCACCCAGT 2639
QY 2690 ATACAGAGAGAGAAACATGACGACATTTATGCTGTGCTCCCGCCCAACAACAACCA 2749
DB 2640 GTGTACG-CCTGCAAGGCTGAGAGACTGACCGCTGACCTGACACAGCGCC-CCAGAAACATA 2697
QY 2750 GACTGTGACTGATCTCTTAGAGACTCAGAGT 2780
DB 2698 CACTGTGACTGATCTCTTAGAGGCTCCAAAT 2728
RESULT 11
AR081724 3147 bp DNA linear PAT 31-AUG-2000
LOCUS AR081724
DEFINITION Sequence 1 from patent US 5972616.
ACCESSION AR081724
VERSION AR081724.1 GI:10008450
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3147)
AUTHORS O'Brien, T.J. and Tanimoto, H.
TITLE TAG-15: an extracellular serine protease overexpressed in breast
JOURNAL Patent: US 5972616-A 1 26-OCT-1999;
FEATURES
source 1..3147
/organism="Unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 60.6%; Score 1883.2; DB 6; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 45 GATGGAACCGCCAAACCATGAGTACGATCGGGGCGGAGAGGCGGGGCTCTCAG 104
DB 5 GAGCGGCTCTCGGGGTATCCATGAGGAGCGATCGGGCGGAGAGGCGGGGCGGAGG 64
QY 105 GACTTCGGGCGGAGCTCAAGTACAACTCCGGCTTAGAGACATGATGATCTTTGAGAG 164
DB 65 GACTTCGGGCGGAGCTCAAGTACAACTCCGGCTTAGAGACATGATGATCTTTGAGAG 124
QY 165 GGTGTGAGAGTCTCTGCTCGAACAATGCGAAGAAAGTGAAGAGGAGGCGCCAGGCGC 224
DB |||||

DB 125 GCGTGAAGTTCCTGCACTCAACAACTCAAGAAAGTGAAGAAAGCATGGCCCGGCGCG 184
QY 225 TGGGATGATCTGATGAGCACTGCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 284
DB 185 TGGGATGATCTGATGAGCACTGCTTCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 244
QY 285 CTGCTGAGCACTTCATTAATGAGAAATGTCGGGTTCAAAAAGTCTTCAATGGCCATCTG 344
DB 245 CTGCTGAGCACTTCATTAATGAGAAATGTCGGGTTCAAAAAGTCTTCAATGGCCATCTG 304
QY 345 AGGATCAACAAATGATCTTTCTGATGAGTGTGATGAGAACTCCCACTCCACAGATTATC 404
DB 305 AGGATCAACAAATGATCTTTCTGATGAGTGTGATGAGAACTCCCACTCCACAGATTATC 364
QY 405 AGCTGAGCAGCCAGGATGAGAGGCGCTGAAAGCTGTAACATGATGATGATGATGATGAT 464
DB 365 AGCTGAGCAGCCAGGATGAGAGGCGCTGAAAGCTGTAACATGATGATGATGATGATGAT 424
QY 465 GGTCTTACCAAGAAAGTGTGCTGATGCTCTTCAATGAGGAGCTGATGATGATGATGAT 524
DB 425 GGTCTTACCAAGAAAGTGTGCTGATGCTCTTCAATGAGGAGCTGATGATGATGATGAT 484
QY 525 TACTGTGAGAGTTCAGATCCCGCCCACTGAGAGAAAGTGTGATGATGATGATGATGAT 584
DB 485 TACTGTGAGAGTTCAGATCCCGCCCACTGAGAGAAAGTGTGATGATGATGATGATGAT 544
QY 585 GTGAGAGGAGTGTGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 644
DB 545 GAGAGAGGAGTGTGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
QY 645 TGTGTGTGAGCTTCCCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 704
DB 605 TCAGTGTGAGCTTCCCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664
QY 705 AGTTTTCCTGAGATGAG 764
DB 665 AGTTTTCCTGAGATGAG 724
QY 765 AACAGTCCCTTACCCGAG 824
DB 725 GACAGCCCTTACCCGAG 784
QY 825 GTGCTGAGCTCACTTCCGAG 884
DB 785 GTGCTGAGCTCACTTCCGAG 844
QY 885 CTGCTGAGCTCACTTCCGAG 944
DB 845 CTGCTGAGCTCACTTCCGAG 904
QY 945 GGCACCTTTCACCTCTCTACAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1004
DB 905 GGCACCTTTCACCTCTCTACAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 964
QY 1005 AGCTGATTAACATTAATGAG 1064
DB 965 ACATGATTAACATTAATGAG 1024
QY 1065 CCAAGATGAG 1124
DB 1025 CCAAGATGAG 1084
QY 1125 TACTATCCAGGCACTTACCCGCAACATCACTGACATGAGAAATATCAAGTGTCCCAAC 1184
DB 1185 TACTATCCAGGCACTTACCCGCAACATCACTGACATGAGAAATATCAAGTGTCCCAAC 1144
QY 1185 AACCGAAAGTGAAGTGTGCTTCAACTCTTCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 1244
DB 1145 AACCGAAAGTGAAGTGTGCTTCAACTCTTCTGATGAGAGAGAGAGAGAGAGAGAGAGAG 1204
QY 1245 GGTCTCTGACCAAGAGATGATGAGATCAACGGGAGAGAGATGATGATGATGATGATGAT 1304
DB 1205 GGTCTCTGACCAAGAGATGATGAGATCAACGGGAGAGAGATGATGATGATGATGATGAT 1264


```

345 AGGATCAAAATGAGATCTTTCTGATGCGATGAGAACTCCACCTCCACAGATTATC 404
305 AGGATCAAAATGAGATTTTGTGATGCTTACGAAACTCCAACTCCATGTGTTGA 364
405 AGCTGGCCAGCGAGTGAAGAGGCGCTGAAGCTGTGATCAATGAAGTCCCTGTCTG 464
365 AGCTGGCCAGCGAGTGAAGAGGCGCTGAAGCTGTGATCAAGCGGAGTCCCATTCCTG 424
465 GGTCCCTACCAAGAAGTGGCTGTAACTGCTTCAAGTGAAGGCGAGTGTCAATGCTTAC 524
425 GGGCCCTACCAAGAAGTGGCTGTGAAGGCTTCAAGCGGAGTGTCAATGCTTAC 484
525 TACTGTGATGATTTGAGCATCCCCCACTGTGAGGAGAAAGTTATGCGGCGATGAGCT 584
485 TACTGTGATGATTTGAGCATCCCCCACTGTGAGGAGAAAGTTATGCGGCGATGAGCT 544
585 GTGAGAGAGTTGATTAATTTGCACTTGGCACTCCGAGCAAGGCACTGAATCTTCTGTCTACA 644
545 GAGAGAGCGCTATGATGCTGCTGCTCCGCGGCGGCTCCCTGAAGTCTTTGTGATCAC 604
645 TCTGTGATGCTTCCCATTTGACCCCAAGATGCTGAGAGACTGAGCAACAGCTGC 704
605 TCAGTGTGCTTTCCCAAGAGACTCCAAAAGTACAGAGAGCCCAAGAGAGCTGC 664
705 AGTTTTCCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 764
665 AGCTTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 724
765 AACAGTCCCTACCGCGGCGAGTCCGCTGCGAGTGGTCTGCGGCGGAGAGCGCGACTCT 824
725 GACAGCGCTTACCGCGGCGAGTCCGCTGCGAGTGGTCTGCGGCGGAGAGCGCGACTCT 784
825 GTGTGATGCTTACCGCTTCCGAAAGCTTTGATGCTGCTGCTGCTGATGATGATGATGATG 884
785 GTGTGATGCTTACCGCTTCCGAAAGCTTTGATGCTGCTGCTGATGATGATGATGATGATG 844
885 CTGTGATGCTTACCGCTTCCGAAAGCTTTGATGCTGCTGCTGATGATGATGATGATGATG 944
845 CTGTGATGCTTACCGCTTCCGAAAGCTTTGATGCTGCTGCTGATGATGATGATGATGATG 904
945 GGCACCTTCTACCGCTTCCGAAAGCTTTGATGCTGCTGCTGATGATGATGATGATGATGATG 1004
905 GGCACCTTCTACCGCTTCCGAAAGCTTTGATGCTGCTGCTGATGATGATGATGATGATGATG 964
1005 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
965 ACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1024
1065 CCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1124
1025 CTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1084
1125 TACTATCCAGGCACTACCCGCGCAATCACTGACATGATGATGATGATGATGATGATGATGATG 1184
1085 TACTATCCAGGCACTACCCGCGCAATCACTGACATGATGATGATGATGATGATGATGATGATG 1144
1185 AACCGGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1244
1145 AACCGGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1204
1245 GGTCTCTGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1304
1205 GGTCTCTGACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1264
1305 CAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1364
1265 CAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1324
1365 TACACGAGACCGGCTTCTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1424
1325 TACACGAGACCGGCTTCTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1384

```

```

QY 1425 GGGATGTTCAATGAGAGAGTGGAGTGGATGCTGCAAGAGAACTGCGCTGCCAGCGCTGG 1484
DB 1385 GGGATGTTCAATGAGAGAGTGGAGTGGATGCTGCAAGAGAACTGCGCTGCCAGCGCTGG 1444
QY 1485 GCGAGTCCCGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1544
DB 1445 GCGAGTCCCGGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1504
QY 1545 AGCTGCAAAAACAGTTTCTGCAAGCGCTTCTGAGTCTGTGATGATGATGATGATGATGATG 1604
DB 1505 AGCTGCAAAAACAGTTTCTGCAAGCGCTTCTGAGTCTGTGATGATGATGATGATGATGATG 1564
QY 1605 GGGAGCGAAGTGAAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1664
DB 1565 GGGAGCGAAGTGAAGAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1624
QY 1665 GGGAGTGTCTCCCTGAGAGCGCAAGTGTATGAGAGAGCACTGTGAGATGAGTGTCT 1724
DB 1625 GGGAGTGTCTCCCTGAGAGCGCAAGTGTATGAGAGAGCACTGTGAGATGAGTGTCT 1684
QY 1725 GACGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1784
DB 1685 GACGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1744
QY 1785 CAAATGAGCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1844
DB 1745 CTAATGAGCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1804
QY 1845 GATGAGCTCCGATGAGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1904
DB 1805 GACGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1864
QY 1905 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1964
DB 1865 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1924
QY 1965 CTGGGCGAAGGCGCACTTGTGTGAGGCTTGTGATGATGATGATGATGATGATGATGATG 2024
DB 1925 CTGGGCGAAGGCGCACTTGTGTGAGGCTTGTGATGATGATGATGATGATGATGATGATG 1984
QY 2025 GCTCATGCTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2084
DB 1985 GCTCATGCTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2044
QY 2085 TTCTGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2144
DB 2045 TTCTGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2104
QY 2145 AAAGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2204
DB 2105 AAAGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2164
QY 2205 CTGAGTCTGAGAGAGTGGTGAAGTACAGACCGTGTGAGCGCCCATGCTGCTGTAT 2264
DB 2165 CTGAGTCTGAGAGAGTGGTGAAGTACAGACCGTGTGAGCGCCCATGCTGCTGTAT 2224
QY 2265 GCTACCATGCTTCTCTGCTGAGAGGCACTGTGAGTCAAGGCTTGGGGGAGAGCAAA 2324
DB 2225 GCTACCATGCTTCTCTGCTGAGAGGCACTGTGAGTCAAGGCTTGGGGGAGAGCAAA 2284
QY 2325 GAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2384
DB 2285 TATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2344
QY 2385 ACCTGAGAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2444
DB 2345 ACCTGAGAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2404
QY 2445 AGTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2504
DB 2405 AGTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2464
QY 2505 GATGGCGAATGTTTCAAGCTGTGTGTGATGATGATGATGATGATGATGATGATGATGATG 2564

```


Db 2465 GATGGGCGGATCTTCCAGGCGGCTGTGGTGGCTGGGAGCGGCTGGCTCAAGAGAAC 2524
 QY 2565 AAGCCAGGCGTGTACAAAGGCTCCCTGTAGTTCGGGACTGGATCAAAAGACACTGGG 2624
 Db 2525 AAGCCAGGCGTGTACAAAGGCTCCCTGTAGTTCGGGACTGGATCAAAAGAGAACACTGGG 2584
 QY 2625 GTATAGAGAGATGAGACAGACGCGGACCAAAACCCACAGGGATGCCGACATGACA 2684
 Db 2585 GTATAGGCGCGGG--GCCAACCAGATGTGTACACCTGGCGGGCCACCAATCGTCCACC 2641
 QY 2685 CCGTATACAGAGAGAGAACATGACATTTATGTGTGGCTCCCGCCCAACACA 2744
 Db 2642 CCAAGTGTGACG--CTGTGAGGCTGTGAGACTGGACCGCTTGACTGACAGCGCC--CCAGA 2699
 QY 2745 ACCCAGACTGTGAATGCTGATCTTTAGGACTGACAGT 2780
 Db 2700 ACATACACTGTGAATCTCAATCTCCAGGGGCTCCAAAT 2735

RESULT 13
 AR229712/c AR229712 3147 bp RNA linear PAT 20-DEC-2002
 LOCUS AR229712 Sequence 18 from patent US 6451500.
 DEFINITION AR229712
 ACCESSION AR229712
 VERSION AR229712.1 GI:27269442
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3147)
 AUTHORS Leon J.W.
 TITLE Imaging member containing heat switchable carboxylate polymer and method of use
 JOURNAL Patent: US 6451500-A 18 17-SEP-2002;
 FEATURES Location/Qualifiers
 source 1..3147
 /organism="Unknown"
 /mol_type="unassigned RNA"

ORIGIN

Query Match 60.6%; Score 1883.2; DB 6; Length 3147;

Best Local Similarity 81.2%; Pred. No. 0;

Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 45 GATCGGACCGCAAAACCATGGGTAGCAATCGGGCGCGCAAGCGCGAGGGGCTCTCAG 104
 Db 3143 GAGCGGCTCGGGGTACCATGGGAGCGATCGGGCCCGCAAGGGCGGAGGGGCCGGAAG 3084
 QY 105 GACTTCGGGGCGGAGCTCAAGTCAACTCCCGCTAGAGAACATGAAATGGCTTTAGAGAG 164
 Db 3083 GACTTCGGGGCGGAGCTCAAGTCAACTCCCGCAAGAAATGAAATGGCTTTAGAGAGAA 3024
 QY 165 GGTGTGAGATTCTGTGCTCGAAACATGCGAAGAAAGTGAAGAGAGAGAGAGAGAGAGAG 224
 Db 3023 GCGGTGAGATTCTGTGCTCGAAACATGCGAAGAAAGTGAAGAGAGAGAGAGAGAGAGAG 2964
 QY 225 TGGGTGTGCTGTGTGAGAGTGTGTCACTTCTTGTCTCTCCCTCAATGAGCTGGCTTG 284
 Db 2963 TGGGTGTGCTGTGTGAGAGTGTGTCACTTCTTGTCTCTCCCTCAATGAGCTGGCTTG 2904
 QY 285 CTGGTGTGAGCACTTCCATTTATGGAATGTGCGGGTTCAAAAAGTCTTCAATGAGCAATCG 344
 Db 2903 CTGGTGTGAGCACTTCCATTTATGGAATGTGCGGGAGTGTGCAATGAGCTTCAATGAGCAATG 2844
 QY 345 AGGATCACAAAATGAGATCTTTCTGTGATGCTATGAGAACTCCAGCTCCAGAGATTATAC 404
 Db 2843 AGGATCACAAAATGAGATTTTGTGTGATGCTATGAGAACTCCAGCTCCAGAGATTATAC 2784
 QY 405 AGCTGTGCGAGCGAGTGAAGAGCGGTGAAGTGTGTGATCAATGAAGTCCCTGTCTCG 464
 Db 2783 AGCTGTGCGAGCGAGTGAAGAGCGGTGAAGTGTGTGATCAAGTGTGATGAGTCCCATTTCTG 2724

QY 465 GGTCTTACCAAGAAAGTGGCTGTATCTGCTTCACTGAGAGGCACTGATCGCTAC 524
 Db 2723 GGCCCTACCAAGAAAGTGGCTGTATCTGCTTCACTGAGAGGCACTGATCGCTAC 2664
 QY 525 TACTGTCAAGATTACAGATTCCTCCACACCTGTGAGAGAGATTGATCGCGCATGCT 584
 Db 2663 TACTGTCAAGATTACAGATTCCTCCACACCTGTGAGAGAGATTGATCGCGCATGCT 2604
 QY 585 GTGAGCGAGTTGAAATGCAATGCGACCCCGAGACAGGGGACTGAAATCTTGTCTTACA 644
 Db 2603 GAGAGCGGATGATGATGCTGCTCCCGGGCGGCTCTGAGTCTTGTGTGATACC 2544
 QY 645 TCTGTGTGCTTCCCATTTGACCCAGATGTGTGAGAGAGACTGAGACCAAGCTGC 704
 Db 2543 TCAGTGTGCTTTCCTCCAGAGACTCCAAACATGACAGAGAGCCAGACAGAGCTGC 2484
 QY 705 AGTTTGTGCTGTATGCTGACAGAGTGAACAGCTTCACTACCTGTGCTTCCC 764
 Db 2483 AGCTTGTGCTGTATGCTGACAGAGTGAACAGCTTCACTACCTGTGCTTCCC 2424
 QY 765 AACAGTCCCTACCGGGGAGATGCGCGCTGTGATGCTGCTGTGAGAGAGAGAGAGAGAG 824
 Db 2423 GACAGCCCTACCGCGCTATGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2364
 QY 825 GTGCTGAGCTCACTTCCGAAAGCTTGAATGTGCTCCTGTGATGAGCATGAGAGAGAG 884
 Db 2363 GTGCTGAGCTCACTTCCGAGCTTGAATGTGCTCCTGTGATGAGCATGAGAGAGAGAG 2304
 QY 885 CTGTGACCGTGTATATAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 944
 Db 2303 CTGTGACCGTGTATATAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2244
 QY 945 GGCACCTTCAACCCGCTCAACAGCTGATCTTCTCTCCAGAAAGCTTCTTGTTC 1004
 Db 2243 GGCACCTTCAACCCGCTCAACAGCTGATCTTCTCTCCAGAAAGCTTCTTGTTC 2184
 QY 1005 AGCTGTATACCAATATAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1064
 Db 2183 ACACGTATACCAATATAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2124
 QY 1065 CCCAAGATAGAGAGCTGTGCGGCTTTTGAAGTACACCCAAAGGACATTTAGCAGAGAG 1124
 Db 2123 CCTAGATAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2064
 QY 1125 TACTATCCAGGACATCCCGGCGCAATCACTGACATGAGATATCAAGTGTCCCAAC 1184
 Db 2063 TACTATCCAGGACATCCCGGCGCAATCACTGACATGAGATATCAAGTGTCCCAAC 2004
 QY 1185 AACCGAAAGTGAAGTGTGCTTCAAACTTCTTATGTGTGAGAGAGAGAGAGAGAGAGAG 1244
 Db 2003 AACCGAAAGTGAAGTGTGCTTCAAACTTCTTATGTGTGAGAGAGAGAGAGAGAGAGAG 1944
 QY 1245 GGTCTCTGACCAAGAGACTATGTGAGATCAACGAGAGAGAGAGAGAGAGAGAGAGAG 1304
 Db 1943 GGTCTCTGACCAAGAGACTATGTGAGATCAACGAGAGAGAGAGAGAGAGAGAGAGAG 1884
 QY 1305 CAGTTGTGTGAG 1364
 Db 1883 CAGTTGTGTGAG 1824
 QY 1365 TACAGGAGACCGGCTTCTAGTGTGAGATCTCTTCAAGTCAAGAGAGAGAGAGAGAGAG 1424
 Db 1823 TACAGGAGACCGGCTTCTAGTGTGAGATCTCTTCAAGTCAAGAGAGAGAGAGAGAGAG 1764
 QY 1425 GGGATGTTCATGTGAG 1484
 Db 1763 GGGATGTTCATGTGAG 1704
 QY 1485 GCAAGCTGCGCGAGATTAAGAGTGAAGGTTAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 1544
 Db 1703 GCAAGCTGCGCGAGATTAAGAGTGAAGGTTAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 1644
 QY 1545 AGGTGCAAAAACAGTTCTGCAAGCCCTTCTGTGTGTGACAGTGTCAAGAGAGAGAG 1604

Db	1643	ACGTGCAAGAACAAAGTTCTGCAAGCCCTCTTCTGAGTCTGGCAAGTGTGAACGACTGC	1584
QY	1605	GGGAGACGAAGTGAACGAGAGAGGAGCTGCAGCTGTCTCTGAGAGTTTCAAGTGTTCAAAT	1664
Db	1583	GGAGACAAACAGCAGCAGCAGAGGGGTGCAAGTTGTCTCGGCCAGACCTTCAGTGTTCAAAT	1524
QY	1665	GGGAAGTGTCTCCTCTCAGAGCCAGAGAGTGTAAATGGGAAGGACAACTGTGGAGATGTGGTCT	1724
Db	1523	GGGAAGTGTCTCTGCAAAAGCCAGCACTGTGCAATGGGAAGGACCACTGTGGAGCGAGTCTCC	1464
QY	1725	GACGAGCTTCATGTGACAGGCGTAATGTGGTCTCTTGACCAAAATATACCTTACCCGCTGG	1784
Db	1463	GACAGAGCTCTCTGCCCCCAAGGTGAAGTGTGTCACTTTGTACCAAAACAACTTACCGCTGG	1404
QY	1785	CAAAATGAGCCTCTGTCTGAGCAGAGGAGCAACCTTGATGTGATGGGAAGACGGACTGTAGC	1844
Db	1403	CTCAATGGGCTCTGTCTTGAAGCAAGGCAACCTTGATGTGACGGGAAGGAGACTGTAGC	1344
QY	1845	GATGGCTCCGATGAGAAAACCTGTGACTGTGGCTGGGCTGGGATCTTTTACCAACAGGCTCGG	1904
Db	1343	GACGAGCTCAGATGAGGAAGACTGTGGACTGTGGGCTGGGCTATTCAAGAAACAGGCTCGT	1284
QY	1905	GTGTGTGGTGGCAAGAAATGGGAGCAGAGGGAGATGGGCTCTGGCAGGTGAACCTTCCAGCC	1964
Db	1283	GTGTGTGGGAGCAGAGATGCGGAATGAGGGCGAGTGGCTCTGGCAGGTGAATGCTGCAATGCT	1224
QY	1965	CTGGGAGCAGGAGCACTTGTGTGGAGCCCTCGCTCATCTCTCTGACTGGCTGTGTCTTGCA	2024
Db	1223	CTGGGAGCAGGAGCACTTGTGGAGTCTTCTCTCATCTCTCTTCACTTCTCCAACTGAGTGTCTGGC	1164
QY	2025	GCTCATGCTCTTTCAGAGATGACAAAATTTCAAGTACTACAGTCAACAGATGTGACAGGCG	2084
Db	1163	GCACACTCTCATGATGATGACAGAGATTTCAAGTACTCAGACCCACAGCAAGTGAACGGCC	1104
QY	2085	TTCTCGGGTCTGCTGGACACAGAGCAAGGCGAGTGGCTCTGGGAGTGCAGAGAGCTAACTCTC	2144
Db	1103	TTCTCGGGCTTGCAGACACAGGCGAGGCGAGCGCCCTTGGGAGTGTGCAGAGAGCAGGCTC	1044
QY	2145	AAACGTATCATCACCACCCCTTCTCTTCAATGATTTTCACTTTCGACTATGACTATGACTGCTTG	2204
Db	1043	AAAGCCATCATCTCCCAACCCCTTCTTCAATGACTTCACTTTCGACTATGACTATGACTGCTTG	984
QY	2205	CTGAGCTGAGAAAGTCCGTGGAGTACACGACCGTCTGGTGGCCCTCATCTGCTTCCCTGAT	2264
Db	983	CTGAGCTGAGAAACCCGACAGAGTACAGCTTCAATGGTGCAGGCCCATTTGCTTCCGAGAC	924
QY	2265	GCTACCCATGTCTTCTCCCTGCTGGCAAGGCAATCTGGGTCAACAGGCTGGGGGACACAAAA	2324
Db	923	GCTTCCCATGTCTTCTCCCTGCTGGCAAGGCAATCTGGGTCAACAGGCTGGGGGACACACCCAG	864
QY	2325	GAGGAGAGTACCGGAGCGCTGATCTCTGCAAGAGGAGTATCCGTATCAACCAAGACC	2384
Db	863	TATGAGAGCACTGGGCGCGCTATCTCTGCAAAAGGATGATCCGGTCAATCAACCAAGACC	804
QY	2385	ACCTGTAGAGACCTTATGCTCCGCAAGGATGACTGTGGAGGCGCCCTTGTCAAGCGCGAGAAA	2444
Db	803	ACCTGTAGAAACCTCTCTGCGCAGAGCAAGTCAAGCGCGCATGATGTGGCTTGTCTCTC	744
QY	2445	AATGGGGGTGTGACCTCCGCAAGGATGACTGTGGAGGCGCCCTTGTCAAGCGCGAGAAA	2504
Db	743	AAGGGGGGTGTGACTTCTGCAAGGATGATTTCCGGGGACCCCTTGTCAAGCGTGGAGGG	684
QY	2505	GATGGGCGAATGTTCCAGGCTGGTGTGTGAGCTGGGGTGAAGAGCTGCGCTCAGAGAAC	2564
Db	683	GATGGGCGAATCTTCCAGGCGCGTGTGTGAGCTGGGGAGAGCGGCTGCGCTCAGAGAAC	624
QY	2565	AAAGCAGGCGTGTACCAAGGCTCCCTGTATGTTCCGGACTGGATCAAAAGACCACTGGG	2624
Db	623	AAAGCAGGCGTGTACCAAGGCTCCCTGTATGTTCCGGACTGGATCAAAAGAACCACTGGG	564
QY	2625	GTTATACAGTATGACAGACGCCCAACCAAAACATCCCAAGGAGTGGCCGACATGACA	2684

Db	563	GTATAGAGGGCCGGG---	GCCACCCCAATGTGTACACTGCGGGGCCACCACTGCTCACC	507
OY	2685	CCTGTGATACAGGAGAGGAACACTGAGACACATTATGCTGTGGCCCTCCCAACACA	2744	
Db	506	CCAGTGTGACG-CCTGCAGCGTGAACACTGAGCCGTGACTGCACACGCGCC-CCAGA	449	
OY	2745	ACCAGACTGTGAATGACTGATCCTTAGGACTCAGAGT	2780	
Db	448	ACATACACTGTGAACCTCATCTCCAGGGCTCCAAAT	413	
RESULT 14				
AR430930		3147 bp	DNA	linear
LOCUS				PAT 18-DEC-2003
DEFINITION				
Sequence 1 from patent US 6649741.				
AR430930				
VERSION				
AR430930.1		GI:40192672		
KEYWORDS				
SOURCE				
ORGANISM				
Unknown.				
Unclassified.				
REFERENCE				
1 (bases 1 to 3147)				
O'Brien,T.J. and Tanimoto,H.				
TAG-15: an extracellular serine protease overexpressed in				
carcinomas				
Patent: US 6649741-A 1 18-NOV-2003;				
Location/Qualifiers				
1. 3147				
/organism="unknown"				
/mol_type="genomic DNA"				
ORIGIN				
Query Match		60.6%; Score 1883.2; DB 6; Length 3147;		
Beet Local Similarity		81.2%; Pred. No. 0;		
Matches 223; Conservative		0; Mismatches 508; Indels 5; Gaps 3		
OY	45	GATCGGACCGGCCAAAACCATGGGTAGACATCGGGCCGCCAAGCCGAGGGGCTCTCAG	104	
Db	5	GAGCGGCTCTGGGGGTACCATGGGAGGAGATCGGGCCGCCAAGGCGGAGGGGCCGAAG	64	
OY	105	GACTTCGGCGGGGACTCAATGATCACTCCCGCTAGAGAAATGAATGGCTTTAGAGG	164	
Db	65	GACTTCGGCGGGGACTCAATGATCACTCCCGCGAGAAAGTGAATGGCTTTAGAGGAA	124	
OY	165	GGTGTGAGGTTCTCTCGTCCGAACATGCGCAAGAGTGAAGACGAGGCCCAAGCGC	224	
Db	125	GGCGTGAAGTTCTCTCGCACTCAACAACGTCAGAAAGTGAAGACATGGCCCGGGCGC	184	
OY	225	TGGGTGTGCTGTGTGAGGAGGAGGCTGTTCAGCTTCTCTTGTCTCTCCTCATGGCTTGG	284	
Db	185	TGGGTGTGTGTGTGAGGAGGAGGCTGTTCAGCTTCTCTTGTCTCTCCTCATGGCTTGG	244	
OY	285	CTGTGTGTGCACTTCTCATTTATGGAATGTGGCGGTTCAAAAAGTCTTCAATGGCCATCG	344	
Db	245	CTGTGTGTGCACTTCTCATTTATGGAATGTGGCGGTTCAAAAAGTCTTCAATGGCCATCG	304	
OY	345	AGGATCACAATGAGATCTTTCTGATGTGATAGAACTCCACTCCACAGATTATC	404	
Db	305	AGGATCACAATGAGATCTTTCTGATGTGATAGAACTCCACTCCACAGATTATC	364	
OY	405	AGCTTGGCCAGCCAGGTGAAGAGGCGCTGAAGCTGTGTACATGAATGATCCTGTCTCG	464	
Db	365	AGCTTGGCCAGCCAGGTGAAGAGGCGCTGAAGCTGTGTACAGCGAGTCCATTCTCG	424	
OY	465	GGTCCCTTACCAAGAAGTCGCTGTATCTGCTTCAAGTGAAGGCGAGTGCATCGCTAC	524	
Db	425	GGTCCCTTACCAAGAAGTCGCTGTATCTGCTTCAAGTGAAGGCGAGTGCATCGCTAC	484	
OY	525	TACTGTGAGTTGAGATCCCGCCACCTGTGTGAGAGAGGCTGATCGCCATGCT	584	
Db	485	TACTGTGAGTTGAGATCCCGCCACCTGTGTGAGAGAGGCTGATCGCCATGCT	544	
OY	585	GTGGAACGAGTTTAACTTGCACCCCGAGCACGGGCACTGAATCTTGTGTCTAC	644	

RESULT 15
 AR430938/c 3147 bp RNA 1linear PAT 18-DEC-2003
 LOCUS AR430938 Sequence 18 from patent US 6649741.
 DEFINITION AR430938
 ACCESSION AR430938
 VERSION AR430938.1 GI:40192680
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1. (bases 1 to 3147)
 O'Brien, T.J. and Taniuchi, H.
 TADG-15: an extracellular serine protease overexpressed in
 carcinomas
 JOURNAL Patent: US 6649741-A 18 NOV-2003;
 FEATURES Location/Qualifiers
 source 1..3147
 /organism="unknown"
 /mol_type="unassigned RNA"
 ORIGIN
 Query Match 60.6%; Score 1883.2; DB 6; Length 3147;
 Best Local Similarity 81.2%; Pred. No. 0;
 Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;
 45 GATCGACCCGCAAAACCATGGTAGCAATCGGGCGCGCAAGCGCGAGGGGCTCTCAG 104
 3143 GAGCGGCTCGGGGTACCATGGGAGCGATCGGGCGCGCAAGGGCGGCGCCGCAAG 3084
 105 GACTTCGGCGCGGAGCTCAAGTACACTCCCGGTAGAGAACATGAGTTGAGAGAG 164
 3083 GACTTCGGCGCGGAGCTCAAGTACACTCCCGGTAGAGAACATGAGTTGAGAGAG 3024
 165 GGTTGAGAGTTCTGCTGCGACATGCGCAAGAAAGTGAAGGCGAGCGCGC 224
 3023 GCGGTGAGTTCTGCTGCGACATGCGCAAGAAAGTGAAGGCGAGCGCGC 2964
 225 TGGGTGAGTTCTGCTGCGACATGCGCAAGAAAGTGAAGGCGAGCGCGC 284
 2663 TGGGTGAGTTCTGCTGCGACATGCGCAAGAAAGTGAAGGCGAGCGCGC 2904
 285 CTGTTGAGTTCTGCTGCGACATGCGCAAGAAAGTGAAGGCGAGCGCGC 344
 2903 CTGTTGAGTTCTGCTGCGACATGCGCAAGAAAGTGAAGGCGAGCGCGC 2844
 345 AGGATCACAATGAGATCTTCTGATGCGATGAGAACTCCACTCCAGAGTTATC 404
 2843 AGGATCACAATGAGATCTTCTGATGCGATGAGAACTCCACTCCAGAGTTATC 2784
 405 AGCTGCGCAGCGAGTGAAGAGCGGCTGAAGTCTGTCATGAGAGTCCCTGCTG 464
 2783 AGCTGCGCAGCGAGTGAAGAGCGGCTGAAGTCTGTCATGAGAGTCCCTGCTG 2724
 465 GGTCCCTACCAAGAAAGTGGCTGTAAGTCTGCTGAGAGGCGAGTGTATCGCTAC 524
 2123 GGTCCCTACCAAGAAAGTGGCTGTAAGTCTGCTGAGAGGCGAGTGTATCGCTAC 2664
 525 TACTGTCAGAGTTGAGATCCCGCAGACCTGAGAGAGAGTTGATGCGCATGCT 584
 2663 TACTGTCAGAGTTGAGATCCCGCAGACCTGAGAGAGAGTTGATGCGCATGCT 2604
 585 GTGAGCGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 644
 2603 GAGGAGCGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2544
 645 TGTGAGTGGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 704
 2543 TGTGAGTGGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2484
 705 AGTTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764

Db 2483 AGCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2424
 Qy 765 AACAGTCCCTACCCGCGAGTGGCGGCTGCGAGTGGCTGCGAGGAGAGCGCGATCT 824
 Db 2423 GACAGCCCTTACCCGCGAGTGGCGGCTGCGAGTGGCTGCGAGGAGAGCGCGATCT 2364
 Qy 825 GTGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 884
 Db 2363 GTGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2304
 Qy 885 CTGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 944
 Db 2303 CTGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2244
 Qy 945 GGCACCTTCTACCCCTCTTACCACTGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1004
 Db 2243 GGCACCTTCTACCCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2184
 Qy 1005 AGCTGATTAACCAATCTGACCGGCGACATCTGCTTGTGAGGCGATTTTCTTCAAG 1064
 Db 2183 ACATGATTAACCAATCTGACCGGCGACATCTGCTTGTGAGGCGATTTTCTTCAAG 2124
 Qy 1065 CCAGATGAGAGCTGCTGCGGCTTTTGTGATGATGATGATGATGATGATGATGATGAT 1124
 Db 2123 CTGAGATGAGAGCTGCTGCGGCTTTTGTGATGATGATGATGATGATGATGATGATGAT 2064
 Qy 1125 TACTATCCAGAGCTTACCCGCGACATGATGATGATGATGATGATGATGATGATGATGAT 1184
 Db 2063 TACTATCCAGAGCTTACCCGCGACATGATGATGATGATGATGATGATGATGATGATGAT 2004
 Qy 1185 AACCGAAGCTGAGAGTGGCTTCAACTCTTCTATGTTGAGTGGCGGCGGAGTGGCG 1244
 Db 2003 AACCGAAGCTGAGAGTGGCTTCAACTCTTCTATGTTGAGTGGCGGCGGAGTGGCG 1944
 Qy 1245 GGTCTGCGACCAAGAGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1304
 Db 1943 GGTCTGCGACCAAGAGCTATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1884
 Qy 1305 CAGTTTGTGAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCA 1824
 Db 1883 CAGTTTGTGAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCAAGAGCA 1764
 Qy 1365 TACACGAGACCGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1424
 Db 1823 TACACGAGACCGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1764
 Qy 1425 GGGATGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1484
 Db 1763 GGGATGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1704
 Qy 1485 GCAGACTGCGCGAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1544
 Db 1703 GCAGACTGCGCGAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1644
 Qy 1545 ACCTGCAAAAACAGTTTCTGAGAGCGCTTCTGAGGCTGATGATGATGATGATGATGAT 1604
 Db 1643 ACCTGCAAAAACAGTTTCTGAGAGCGCTTCTGAGGCTGATGATGATGATGATGATGAT 1584
 Qy 1605 GGGGACGAGAGTGAAGAGAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1664
 Db 1583 GGGGACGAGAGTGAAGAGAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1524
 Qy 1665 GGGAGTGTCTTCTCTCAGAGCCAGAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAG 1724
 Db 1523 GGGAGTGTCTTCTCTCAGAGCCAGAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAG 1664
 Qy 1725 GACGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1784
 Db 1463 GACGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1404
 Qy 1785 CAAAATGAGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1844
 Db 1403 CTAATGAGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1344

BLANK
PAGE